

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 04:38:43 ; Search time 18292 Seconds  
(without alignments)  
11663.477 Million cell updates/sec

Title: US-09-837-138-1  
Perfect score: 4403  
Sequence: 1 ttccgcacgagcgcggttg.....accgcgttgagctccagct 4403

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sta.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4164.2	94.6	4406	9 AF058696	AF058696 Homo sapi
2	4123.4	93.6	4412	9 AF051334	AF051334 Homo sapi
3	4115.2	93.5	4388	6 CQ725600	CQ725600 Sequence
4	4108.4	93.3	4386	6 AR233599	AR233599 Sequence
5	4102	93.2	4466	9 BC071590	BC071590 Homo sapi
6	4038.4	91.7	4423	11 BV177869	BV177869 sqnm97272
7	1862.8	42.3	55134	9 AY566246	AY566246 Homo sapi
8	1862.8	42.3	56500	9 AB013139	AB013139 Homo sapi
9	1862.8	42.3	184919	9 AF069291	AF069291 Homo sapi
10	1862.8	42.3	331864	9 AF049895	AF049895 Homo sapi
11	1861.2	42.3	107549	9 AC074178	AC074178 Homo sapi
12	1836.4	41.7	2044	6 BD155805	BD155805 Primer fo
13	1836.4	41.7	2044	6 AX875862	AX875862 Sequence
14	1836.4	41.7	2044	9 AK001017	AK001017 Homo sapi
15	1416.4	32.2	2605	10 AF218575	AF218575 Rattus no
16	1405	31.9	2503	10 BC055061	BC055061 Mus muscu
17	1400.4	31.8	2484	10 AF076687	AF076687 Mus muscu
18	1398.8	31.8	2473	10 AF092840	AF092840 Mus muscu
19	1398.6	31.8	2542	10 BC044773	BC044773 Mus muscu

20	1395.6	31.7	2412	10	AB016988	Mus muscu
21	1391	31.6	2518	10	BC003719	Mus muscu
22	694	15.8	72910	2	AC023998	Homo sapi
23	684	15.5	2533	5	AF230342	Gallus ga
24	671	15.2	752	6	BD145652	Primer fo
25	671	15.2	752	6	AX865590	Sequence
26	564.6	12.8	622	6	BD179056	Cancer as
27	552.2	12.5	646	6	BD265337	Compounds
28	552.2	12.5	646	6	AR401323	Sequence
29	552.2	12.5	646	6	AX192776	Sequence
30	538.6	12.2	544	6	AX341490	Sequence
31	520.4	11.8	543	6	BD151180	Primer fo
32	520.4	11.8	543	6	AX871118	Sequence
33	478.4	10.9	481	6	CQ681271	Sequence
34	456	10.4	2444	5	AY207367	Xenopus l
35	455.6	10.3	2292	5	AV312176	Xenopus l
36	438.2	10.0	439	6	CQ662253	Sequence
37	422.4	9.6	614	11	BV174416	sqnm71594
38	358.4	8.1	396	6	AX067375	Sequence
39	354.2	8.0	1811	5	CR848449	Xenopus t
40	349.6	7.9	1666	5	BC082430	Xenopus l
41	275	6.2	72910	2	AC023998	Homo sapi
42	273.4	6.2	434	11	G66922	csnphb1-pc
43	264.4	6.0	421	11	G30107	human STS S
44	212.6	4.8	475	6	AX401127	Sequence
45	210.8	4.8	238844	2	AC108261	Rattus no

ALIGNMENTS

RESULT 1  
AF058696  
LOCUS AF058696 4406 bp mRNA linear PRI 22-MAY-2003  
DEFINITION Homo sapiens cell cycle regulatory protein p95 (NBS1) mRNA,  
Complete cds.  
ACCESSION AF058696  
VERSION AF058696.2 GI:30995499  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 4406)  
AUTHORS Carney, J.P., Maser, R.S., Olivares, H., Davis, E.M., Le Beau, M.,  
Yates, J.R., III, Hays, L., Morgan, W.F. and Petrini, J.H.  
TITLE The hMre11/hRad50 protein complex and Nijmegen breakage syndrome:  
linkage of double-strand break repair to the cellular DNA damage  
response  
JOURNAL Cell 93 (3), 477-486 (1998)  
MEDLINE 98250063  
PUBMED 9590181  
REFERENCE 2 (bases 1 to 4406)  
AUTHORS Carney, J.P., Maser, R.S., Olivares, H.A., Davis, E.M., Le Beau, M.,  
Yates, J.R., III, Hays, L., Morgan, W.F. and Petrini, J.H.J.  
TITLE Direct Submission  
JOURNAL Submitted (08-APR-1998) Genetics, University of Wisconsin, 445  
Henry Mall, Madison, WI 53706, USA  
COMMENT On May 22, 2003 this sequence version replaced gi:3098674.  
Sequence updated by GenBank staff.  
FEATURES  
Location/Qualifiers  
source  
1..4406  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8q21.3"  
1..4406  
/gene="NBS1"  
53..2317  
/gene="NBS1"  
/function="DNA repair"  
/note="nibrin"

/codon\_start=1  
/product="cell cycle regulatory protein p95"  
/protein\_id="AAC39752.1"  
/db\_xref="GI:3098675"  
/translation="MKKLIFPAAGPAGGEPYRLLTGVEYVVRKNCAILIENDQSISRN  
HAYLTANFSTNLISQDEIPVLTKDNISKYGTFTVNEKMQNGFSRLTKSGDGLITFGVP  
GSKFRIEYPLVACSSCLDVSQGTALNOAILQGGFTVNNWTEECTHLVMVSVKVTIK  
TICALICGRPIVKPEYFTEFLKAVQSKOPPOLSFYPLDPSIGSKNVDLSGRQER  
KQIFPKGTPIFLNAKHKKLSSAVFGGGEARLITEENSEEHNFFLAPGTCTVVDGIT  
NSQTLIPDCQKWISIMDLQROGLRPIPEAEIGLAVIFMTTKNTCDQGHPSLGLK  
TTTTGPSLQGVSDSKLMPSPAPVNTTYVADTESQADTWLDSERPKEIKVSKMEQK  
FRLMSQADPTKESCTSSNNMSWNTLAKMRI PNYSLSPTKLPISINKSKDRASQQ  
QTNISINYFQPSFKTERBENQEMSSKSAR IETSCLLEOTOPATPSLWKNEOHL  
SENEPVDTHSDNNLFTDLDLKSIVKNSASKSHAAEKLRSNKREMDVIAIDEVLEQL  
FKDTELEIDVVKQBEDVNVVRKPRMDIETNDFDSBEAVPESKISQENEIGKKR  
ELKESLSAKEIQNDKLQDSEMLPKLLLTPEFRSLVINKSTSNRPSGINDDYGLQ  
KNPKFKFTYTPQAGLPHIIGGSLIAHHARKNTLBEWLREQMEVQNHQKEBSLA  
DDLFRNPNYLKRRR"

ORIGIN

Query Match 94.6%; Score 4164.2; DB 9; Length 4406;

Best Local Similarity 97.6%; Pred. No. 0;

Matches 4303; Conservative 7; Mismatches 10; Indels 87; Gaps 5;

QY	1	TTCCGACAGGCGCGGTTCACGTGCGCCCGCCAGCCCTGAGGAGCCGCGCATGTGGAA	60
DB	1	TTCCGACAGGCGCGGTTCACGTGCGCCCGCCAGCCCTGAGGAGCCGCGCATGTGGAA	60
QY	61	ACTGCTGCCCGCGCGCGCGCGGAGGAGAACCATACAGACTTTTGACTGGCGTTGA	120
DB	61	ACTGCTGCCCGCGCGCGCGCGGAGGAGAACCATACAGACTTTTGACTGGCGTTGA	120
QY	121	GTACGTTGTTGGAAGGAAAACHTGTGCCATCTTAATGAAAATGATCAGTCGATGAGCG	180
DB	121	GTACGTTGTTGGAAGGAAAACHTGTGCCATCTTAATGAAAATGATCAGTCGATGAGCG	180
QY	181	AAATCATGCTGTCTAACTGCTAACTTTCTGTAAACCAACTCAGTCAACACATCAAAAT	240
DB	181	AAATCATGCTGTCTAACTGCTAACTTTCTGTAAACCAACTCAGTCAACACATCAAAAT	240
QY	241	CCCTGTATTGACATTAAGATAAATCTAAGTATGGTACCTTTGTTAAATGAGGAAAAAT	300
DB	241	CCCTGTATTGACATTAAGATAAATCTAAGTATGGTACCTTTGTTAAATGAGGAAAAAT	300
QY	301	GCAGATGGCTTTCCGGAACCTTTGAAGTCGGGGATGGTATTAATCTTTGGAGTGTTCG	360
DB	301	GCAGATGGCTTTCCGGAACCTTTGAAGTCGGGGATGGTATTAATCTTTGGAGTGTTCG	360
QY	361	AAGTAAATTCAGAAATAGAGTATGAGCTTTGGTTGTCATGCTCTTTGTTAGATGCTC	420
DB	361	AAGTAAATTCAGAAATAGAGTATGAGCTTTGGTTGTCATGCTCTTTGTTAGATGCTC	420
QY	421	TGGGAAAACCTGCTTTAAATCAAGCTATATGCAACTTGGAGGATTTACTGTAAACAATG	480
DB	421	TGGGAAAACCTGCTTTAAATCAAGCTATATGCAACTTGGAGGATTTACTGTAAACAATG	480
QY	481	GACAGAAATGCACCTCTGTCATGTTGATCAGTGAAGTTACCATTAACAATATG	540
DB	481	GACAGAAATGCACCTCTGTCATGTTGATCAGTGAAGTTACCATTAACAATATG	540
QY	541	TGCACCTCATTTGTTGGAGCTCCAATGTTAAAGCCAGAAATATTTTACTGAAATCTCTGAAAGC	600
DB	541	TGCACCTCATTTGTTGGAGCTCCAATGTTAAAGCCAGAAATATTTTACTGAAATCTCTGAAAGC	600
QY	601	AGTTCAAGTCCAGAAGCAGCTCCCAAAATTTGAAAGTTTTTACCACCTCTTGATGAACC	660
DB	601	AGTTCAAGTCCAGAAGCAGCTCCCAAAATTTGAAAGTTTTTACCACCTCTTGATGAACC	660
QY	661	ATCTATTGGAGTAAATATTTGATCTGTGAGCAGCGGAGGAAAGAAACAAATCTTCAA	720
DB	661	ATCTATTGGAGTAAATATTTGATCTGTGAGCAGCGGAGGAAAGAAACAAATCTTCAA	720
QY	721	AGGGAAAAACATTTATATTTTGAATGCCAAACAGCATAAAGAAATTGAGTTCCGCGATTGT	780

DB	721	AGGGAAAAACATTTATATTTTGAATGCCAAACAGCATAAAGAAATTGAGTTCCGCGATTGT	780
QY	781	CTTTGGAGGTGGGAGCTAGTGTGATTAACAGAGAGAAATGAAGAAACATAAATTTCTT	840
DB	781	CTTTGGAGGTGGGAGCTAGTGTGATTAACAGAGAGAAATGAAGAAACATAAATTTCTT	840
QY	841	TTTGGCTCCGGAAACGTGTGTTGTGATACAGAGAAATACAAACTCACAGACCTTAATTC	900
DB	841	TTTGGCTCCGGAAACGTGTGTTGTGATACAGAGAAATACAAACTCACAGACCTTAATTC	900
QY	901	TGACTGTTCAGAGAAATGGAATTCAGTCAATATGATATGCTCCAAAGGCAAGGCTTAG	960
DB	901	TGACTGTTCAGAGAAATGGAATTCAGTCAATATGATATGCTCCAAAGGCAAGGCTTAG	960
QY	961	ACCTATTTCTGAAGCAGAAATTTGGATTTGGCGGTGATTTTTCATGACTACAAAGAAAT	1020
DB	961	ACCTATTTCTGAAGCAGAAATTTGGATTTGGCGGTGATTTTTCATGACTACAAAGAAAT	1020
QY	1021	TGACTCTCAGGGCCATCCAGGTACAGGATTAAGACAACTCCAGGACCAAGCCTTTC	1080
DB	1021	TGACTCTCAGGGCCATCCAGGTACAGGATTAAGACAACTCCAGGACCAAGCCTTTC	1080
QY	1081	ACAAGCGCTGTGATGATGAAAACTAATGCAAGCGCCCCAGTGAACACTACAACTA	1140
DB	1081	ACAAGCGCTGTGATGATGAAAACTAATGCAAGCGCCCCAGTGAACACTACAACTA	1140
QY	1141	CGTAGCTGACACAGAAATCAGAGCAAGCAGATACATGGGATTTGAGTAAAGGCCAAAGA	1200
DB	1141	CGTAGCTGACACAGAAATCAGAGCAAGCAGATACATGGGATTTGAGTAAAGGCCAAAGA	1200
QY	1201	AATCAAGCTCTCCAAATGGAAACAAATTCAGAAATGCTTTTCAAGAGCGACCCACTGT	1260
DB	1201	AATCAAGCTCTCCAAATGGAAACAAATTCAGAAATGCTTTTCAAGAGCGACCCACTGT	1260
QY	1261	AAAGAGTCTCTCAAAACCAAGCTCTAATAATAATAGTATGTTCAAAATCTTTGGCTAA	1320
DB	1261	AAAGAGTCTCTCAAAACCAAGCTCTAATAATAATAGTATGTTCAAAATCTTTGGCTAA	1320
QY	1321	GATGAAATCCCAACTATCAGCTTTCCAACTAAATTTGCCAAGTATAAATAAAGTAA	1380
DB	1321	GATGAAATCCCAACTATCAGCTTTCCAACTAAATTTGCCAAGTATAAATAAAGTAA	1380
QY	1381	AGATAGGCTCTCAGCAGCAGCAGCAACCTCCATCAGAACTACTTTTCAGCCGCTTAC	1440
DB	1381	AGATAGGCTCTCAGCAGCAGCAGCAACCTCCATCAGAACTACTTTTCAGCCGCTTAC	1440
QY	1441	CAAAAAAGGGAAGGAGTGAAGAAATCAAGAAATGTCTTCATGCAAAATCAGCAAGAT	1500
DB	1441	CAAAAAAGGGAAGGAGTGAAGAAATCAAGAAATGTCTTCATGCAAAATCAGCAAGAT	1500
QY	1501	AGAAACGTCTGTTCTCTTTTAGAACAAACACAACTGTCACCTCATTTGTGGAANA	1560
DB	1501	AGAAACGTCTGTTCTCTTTTAGAACAAACACAACTGTCACCTCATTTGTGGAANA	1560
QY	1561	TAAGAGCAGCATCTATCTGAGATGAGCTGTGACACAACTCAGACAAATACTTAT	1620
DB	1561	TAAGAGCAGCATCTATCTGAGATGAGCTGTGACACAACTCAGACAAATACTTAT	1620
QY	1621	TACAGATACAGATTTTAAATCTATTGTGAAAAATCTGCCAGTAAATCTCATGCTGAGA	1680
DB	1621	TACAGATACAGATTTTAAATCTATTGTGAAAAATCTGCCAGTAAATCTCATGCTGAGA	1680
QY	1681	AAAGCTAAGATCAAAATAAAGGAAATGGATGATGTGGCCATAGAGATGAAGTATT	1740
DB	1681	AAAGCTAAGATCAAAATAAAGGAAATGGATGATGTGGCCATAGAGATGAAGTATT	1740
QY	1741	GGAAACAGTTATTCAAGGACACAAAACAGAGTTTGAAGTTGATGTGAAAGTTCAAAAA	1800
DB	1741	GGAAACAGTTATTCAAGGACACAAAACAGAGTTTGAAGTTGATGTGAAAGTTCAAAAA	1800
QY	1801	GGAGGAAGATGTCAAATGTTTAGAAAAAGGCCAAGGATGGATATAGAAAAACAAATGAC	1860
DB	1801	GGAGGAAGATGTCAAATGTTTAGAAAAAGGCCAAGGATGGATATAGAAAAACAAATGAC	1860

Db 1801 GGAGGAAGATGTCATAGTTTAAAGGCAAGATGGAATATAGAAACAAATGACACTTT 1860  
Qy 1861 CAGTGATGAAGCAGTACAGAAAGTAGCAAAATATCTCAAGAAATGAAATTTGGGAAGAA 1920  
Db 1861 CAGTGATGAAGCAGTACAGAAAGTAGCAAAATATCTCAAGAAATGAAATTTGGGAAGAA 1920  
Qy 1921 ACGTGAACCTCAAGGAAGACTCAGTATGGTCAGCTAAGAAATATCTAAACATGACAAACT 1980  
Db 1921 ACGTGAACCTCAAGGAAGACTCAGTATGGTCAGCTAAGAAATATCTAAACATGACAAACT 1980  
Qy 1981 TCAGGATGATGAGTAGCTTCCAAAAAGCTGTTATTGACCTGAAATTTAGATCACTGGT 2040  
Db 1981 TCAGGATGATGAGTAGCTTCCAAAAAGCTGTTATTGACCTGAAATTTAGATCACTGGT 2040  
Qy 2041 GATTAAAAACTCTACTTCCAGAAATCCGCTCGCATAAAATGATGATTATGGTCAACTAAA 2100  
Db 2041 GATTAAAAACTCTACTTCCAGAAATCCGCTCGCATAAAATGATGATTATGGTCAACTAAA 2100  
Qy 2101 AAATTTCAAGAAATTCAAAAAGTCAATATCTCGAGCGAGGAAACTTCCACACATCAT 2160  
Db 2101 AAATTTCAAGAAATTCAAAAAGTCAATATCTCGAGCGAGGAAACTTCCACACATCAT 2160  
Qy 2161 TGGAGGATCAGATCTAATAGCTCATCATGCTCGAAAGAAATACAGAACTAGAGAGTGGCT 2220  
Db 2161 TGGAGGATCAGATCTAATAGCTCATCATGCTCGAAAGAAATACAGAACTAGAGAGTGGCT 2220  
Qy 2221 AAGCGAGGAAATGGAGGTACAAAATCAACATGCAAAAGAGAGTCTCTTCTGATGATCT 2280  
Db 2221 AAGCGAGGAAATGGAGGTACAAAATCAACATGCAAAAGAGAGTCTCTTCTGATGATCT 2280  
Qy 2281 TTTTAGATCAATTCCTTTATTTAAAAAGGAGAAATAACTGAGGATTTTAAAAAGAGCCCA 2340  
Db 2281 TTTTAGATCAATTCCTTTATTTAAAAAGGAGAAATAACTGAGGATTTTAAAAAGAGCCCA 2340  
Qy 2341 TGGAAAACTTCTAGTAGCATCTACTTCAGGCCCAACAGGTTATATGAATATATAGTG 2400  
Db 2341 TGGAAAACTTCTAGTAGCATCTACTTCAGGCCCAACAGGTTATATGAATATATAGTG 2400  
Qy 2401 TATAGAAGCGATTTAAGTTACAAATGTTTTATGCGCTAAAATTTATTAATAAAATGCACAA 2460  
Db 2401 TATAGAAGCGATTTAAGTTACAAATGTTTTATGCGCTAAAATTTATTAATAAAATGCACAA 2460  
Qy 2461 AACTTTGATTTCTTTGATGAACAAATGTTTGTGTCGTTTTGAGGCTTTGTCATTTGCAT 2520  
Db 2461 AACTTTGATTTCTTTGATGAACAAATGTTTGTGTCGTTTTGAGGCTTTGTCATTTGCAT 2520  
Qy 2521 CTTTTTTTCAATTTTAAATGTTTTGTTTTATTAATAGTTAATATAGTACAGAGTTCAAA 2580  
Db 2521 CTTTTTTTCAATTTTAAATGTTTTGTTTTATTAATAGTTAATATAGTACAGAGTTCAAA 2580  
Qy 2581 ATTCTAAATRTAGCTAAGGTAAAGGACTAAAGTCAACCTTCCACCATTGTCCTAGCTACT 2640  
Db 2581 ATTCTAAATRTAGCTAAGGTAAAGTCAACCTTCCACCATTGTCCTAGCTACT 2640  
Qy 2641 ----- 2640  
Db 2640 TGGTTCCTCCCTCAGAAAAAATTCATGATCTCATTTCTTATGAATCTTTCCAGGGATTTT 2699  
Qy 2641 -----TATTTTAAATAATTTCTCTACACAAATGATAGCATACATA 2681  
Db 2700 TGAGTCCCTATTCAAATTCCTATTTTAAATAATTTCTCTACACAAATGATAGCATACATA 2759  
Qy 2682 TGCAGTGTCTACACCTTGCTTTTCTTTTACTTTAGTAAAGATTAAAAATTTATAGGAATATCAAT 2741  
Db 2760 TGCAGTGTCTACACCTTGCTTTTCTTTTCTTTTCCATTATGCTGCTAGTCTTACCTAACTCTGTGA 2818  
Qy 2742 ATAATGTTTTTAATATTTTTCTTTTCCATTATGCTGCTAGTCTTACCTAACTCTGTGA 2801  
Db 2819 ATAATGTTTTTAATATTTTTCTTTTCCATTATGCTGCTAGTCTTACCTAACTCTGTGA 2878  
Qy 2802 TCCAAACAAATGGCTTCAGTGTGCGAGTGTCAACCTACATGTTATCTAGTACTAGAAA 2861  
Db 2879 TCCAAACAAATGGCTTCAGTGTGCGAGTGTCAACCTACATGTTATCTAGTACTAGAAA 2938

Qy 2862 CTGAAGACCATGTGGAGACTTTCATCAAAACATGGGTTTAGTTTTCCACGAATGGAAGAC 2921  
Db 2939 CTGAAGACCATGTGGAGACTTTCATCAAAACATGGGTTTAGTTTTCCACGAATGGAAGAC 2998  
Qy 2922 CTGTACCCCTTTTGGTGTCTTACTAGCTGGTGGGTGTCTGTGTTTGGAGCTTATTTAG 2981  
Db 2999 CTGTACCCCTTTTGGTGTCTTACTAGCTGGTGGGTGTCTGTGTTTGGAGCTTATTTAG 3058  
Qy 2982 AGTCCTAGTGTCTTACTACTTATAAGTAGAAATGGTGAGATTTGTTTTCTTTCTACCKTA 3041  
Db 3059 AGTCCTAGTGTCTTACTACTTATAAGTAGAAATGGTGAGATTTGTTTTCTTTTCTACCTTA 3118  
Qy 3042 AAGGGAGATGGTAAGAAAACAATGAATGTCTTTTTCAAAATTTTATTTGCAAGTGATTTTC 3101  
Db 3119 AAGGGAGATGGTAAGAAAACAATGAATGTCTTTTTCAAAATTTTATTTGCAAGTGATTTTC 3178  
Qy 3102 AAGTCTGTCTTCAAAAATATATTCATGTACTCTGTGATCCAGCAAGAGGGAGTTCAGTTC 3161  
Db 3179 AAGTCTGTCTTCAAAAATATATTCATGTACTCTGTGATCCAGCAAGAGGGAGTTCAGTTC 3238  
Qy 3162 AAGAGTCACCTACAACTGATTTAGTTGTTAGAGAAATGAGAAATGGAACAGTGAGGAATGGA 3221  
Db 3239 AAGAGTCACCTACAACTGATTTAGTTGTTAGAGAAATGAGAAATGGAACAGTGAGGAATGGA 3298  
Qy 3222 GGCCATATTTTCCATGACTTCCCTTTGTAACAGAGCAACAGAGGCAAGAGGCTGGCC 3281  
Db 3299 GGCCATATTTTCCATGACTTCCCTTTGTAACAGAGCAACAGAGGCAAGAGGCTGGCC 3358  
Qy 3282 TCTACATCACTCCACCTTCCAAAATCTTGTGGAAGTGCATCTACTTGCACAGAACCAATTT 3341  
Db 3359 TCTACATCACTCCACCTTCCAAAATCTTGTGGAAGTGCATCTACTTGCACAGAACCAATTT 3418  
Qy 3342 AACTTACTTCCAAAGTCTCGCTGCTTGCAGGTGGAACCTCAGCTGCAAGGGAGTTAGGGA 3401  
Db 3419 AACTTACTTCCAAAGTCTCGCTGCTTGCAGGTGGAACCTCAGCTGCAAGGGAGTTAGGGA 3478  
Qy 3402 AATGAAGGCTCTTTTTTAAAAAGCTTCTCAGCCCTCTAGGGCAACAGAAATGGGTGAGCC 3461  
Db 3479 AATGAAGGCTCTTTTTTAAAAAGCTTCTCAGCCCTCTAGGGCAACAGAAATGGGTGAGCC 3538  
Qy 3462 AATCTGCAATTTCTACTACAGGCAATGAGACCAAGTTAGATTTGAAATATTTATAGAGAG 3521  
Db 3539 AATCTGCAATTTCTACTACAGGCAATGAGACCAAGTTAGATTTGAAATATTTATAGAGAG 3598  
Qy 3522 TTATGAACACTTTAAATTTATGATAGTGTATGATTCATTTGGATAGAACATGGGATACTTTAGA 3581  
Db 3599 TTATGAACACTTTAAATTTATGATAGTGTATGATTCATTTGGATAGAACATGGGATACTTTAGA 3658  
Qy 3582 AGTAGAATTTGACAGGCAATTTAGTTGATGAAATGGAGTCAATTTGAGTCTTCTTAAATAGCC 3641  
Db 3659 AGTAGAATTTGACAGGCAATTTAGTTGATGAAATGGAGTCAATTTGAGTCTCTTAAATAGCC 3718  
Qy 3642 ATGTATCAATTTACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGG 3701  
Db 3719 ATGTATCAATTTACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGG 3778  
Qy 3702 AATAAATCGACAGATTAATATTTGTTCTGTCATGCCCAACATCCCTTTCTTAAGGAAGA 3761  
Db 3779 AATAAATCGACAGATTAATATTTGTTCTGTCATGCCCAACATCCCTTTCTTAAGGAAGA 3838  
Qy 3762 CTGCCCCCTACTATAGCAGTTTTTATATTTGTCATTTTATGAATATATGAATGAGGAGTTTC 3821  
Db 3839 CTGCCCCCTACTATAGCAGTTTTTATATTTGTCATTTTATGAATATATGAATGAGGAGTTTC 3898  
Qy 3822 TGTGACCTCTGTCTTTTCAAAATATTTGGTGTGTGTCAGTATTTTCCCTTTTAAACCT 3881  
Db 3899 TGTGACCTCTGTCTTTTCAAAATATTTGGTGTGTGTCAGTATTTTCCCTTTTAAACCAT 3958  
Qy 3882 TCCCAATTCGGGTGTAGTGGATGTTTCCATTTTGGGTTTTTAAATTTGTATATCCCTGAT 3941  
Db 3959 T-CCAATCGGTGTAGTGTATG- - - - -TTTCAATTTTGGTTTTTAAATTTGTATATCCCTGAT 4012

QY 3942 AGCTATAATTGGGTCATAGAAATCTTTATACATTTCTAGATCAAGTCTCTTGCGGATA 4001  
|||||  
Db 4013 AGCTATAATTGGGTCATAGAAATCTTTATACATTTCTAGATCAAGTCTCTTGCGGATA 4072  
|||||  
QY 4002 TACGTTATGAGATATTACACCTAGTCTGTGGCTTGAGTGTCTTTTATGCTTTTGTATG 4061  
|||||  
Db 4073 TATGTTATGAGATATTACACCTAGTCTGTGGCTTGAGTGTCTTTTATGCTTTTGTATG 4132  
|||||  
QY 4062 AATAGAAGTTTTAAATTTTGACAAAGTCAAAATTTATTTTCTTTTGTGATATTTT 4121  
|||||  
Db 4133 AATAGAAGTTTTAAATTTTGACAAAGTCAAAATTTATTTTCTTTTGTGATATTTT 4192  
|||||  
QY 4122 TCTCTCAATTTAAACCCCAAGATTTTCAGATATTCTGCTCTATTATATAAACTTTATATTT 4181  
|||||  
Db 4193 TCTCTCAATTTAAACCCCAAGATTTTCAGATATTCTGCTCTATTATATAAACTTTATATTT 4252  
|||||  
QY 4182 TTATATTTGTGATCTACCTTTGAATTTGATATGTTGTGAATTTATGGATCAGGGTCTT 4241  
|||||  
Db 4253 TTATATTTGTGATCTACCTTTGAATTTGATATGTTGTGAATTTATGGATCAGGGTCTT 4312  
|||||  
QY 4242 TTTTCCGCCATACAAAGTATCCAGTCATTTGTAACACTGTTTATTGAAAGAAATATCTTTT 4301  
|||||  
Db 4313 TTTTCCGCCATACAAAGTATCCAGTCATTTGTAACACTGTTTATTGAAAGAAATATCTTTT 4372  
|||||  
QY 4302 CCTCATTAATTAACCTTTGCCAATTAGT 4328  
|||||  
Db 4373 CCTCATTAATTAACCTTTGCCAATTAGT 4399  
|||||

RESULT 2  
AF051334  
LOCUS Homo sapiens nibrin (NBS) mRNA, complete cds. PRI 09-MAY-1998  
DEFINITION AF051334  
ACCESSION AF051334  
VERSION AF051334.1 GI:3126794  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 4412)  
AUTHORS Varon,R., Vissinga,C., Platzter,M., Cerosaletti,K.M.,  
Chrzanowska,K.H., Saar,K., Beckmann,G., Seemanova,E., Cooper,P.R.,  
Nowak,N.J., Stumm,M., Weenase,C.M.R., Gatti,R.A., Wilson,R.K.,  
Digweed,M., Rosenchal,A., Sperling,K., Concannon,P. and Reis,A.  
Nibrin, a novel DNA double-strand break repair protein, is mutated  
in Nijmegen breakage syndrome  
Cell 93 (3), 467-476 (1998)  
JOURNAL MEDLINE  
PUBMED 98250062  
REFERENCE 2 (bases 1 to 4412)  
AUTHORS Varon,R. and Platzter,M.  
Direct Submission  
TITLE Submitted (26-FEB-1998) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr.11, Jena 07745, Germany  
FEATURES  
source  
1..4412  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
gene 1..4412  
/gene="NBS"  
CDS 52..2316  
/gene="NBS"  
/notes="mutated in Nijmegen Breakage syndrome"  
/codon\_start=1  
/evidence=not experimental  
/product="nibrin"  
/protein\_id="AAC39732.1"  
/db\_xref="GI:3126795"  
/translation="MWKLLPFAAGPAGGEPYRLLTGVYVVGRRKNCAILIENDQSTSRN  
HAVLTANFVNLSQDEIPLVTLKDNYSKYGTFFVNEKRMQNGFSRLKSGDGTTFGVF  
GSKFRIEYPLVACSSCLDVSGKLTALNQAILQGGFTVNNNTEECTHLVWVSVKYIK

TICALICGRPIVKPEYFTEFLKAVBSKKOPPOIESFYPPLEBPSIGSKNVDLSGRQER  
KOIPKKTFFILNAKOHKKLASAVVGGGEARLIITEENEEHNFFLAECTCVVDGTGIT  
NSQTLIPQCQKNQIOSIMDLQROGLRPIPEAIEGLAVIPMTKNYCDPQHPSTGLK  
TTTPPSUSQGVNSDEKLPMSAPVNTTIVADTESEQADTWDLSERPKEIKVSKWEQK  
FRMSQDAPTQKESCTSSNNNSMVSNTLAKMRIPNYQSPTKLPISINKSKDRASQQQ  
QTNISIRNTFQPSKTRDERDENOEMSSCKSARIETSCSLLEQTPATPSLWKNKQHL  
SENEPVDNTSNNLFTDLDKSIIVKNSASKHAAEKLRSNKKREMDVVAIEDVLEJQL  
PKDTPELEIDVKQKQBEDVNVKRPRMDIETNDTPSDEAVPESKISOENEIGCKR  
ELKEDLSHSAKEISNNDKLODDSEMLPKLLLTPEPSLVIKNSTSRNPSGINDDYGOL  
KNFKPKVTTTPGAKLPHIIGGSLIIAHARKNTELEBWLKQEMEVONQHAKESLSLA  
DDLFRNYPYLKRRR"  
polyA\_signal 2446..2451  
/gene="NBS"  
/evidence=experimental  
polyA\_signal 4392..4397  
/gene="NBS"  
/evidence=experimental  
ORIGIN  
Query Match 93.6%; Score 4123.4; DB 9; Length 4412;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 4287; Conservative 7; Mismatches 28; Indels 91; Gaps 6;  
QY 13 CGCGGTTGCACGTCGCGCCCGCCAGCCCTGAGGAGCCGACCGATGTGTGAAACTCTGTGCCCC 72  
|||||  
Db 12 CGCGGTTGCACGTCGCGCCCGCCAGCCCTGAGGAGCCGACCGATGTGTGAAACTCTGTGCCCC 71  
|||||  
QY 73 CGCGGCGCGGCGAGGAGAAACCATACAGACTTTTGTGACTGCGGTTGAGTACGTTGTGG 132  
|||||  
Db 72 CGCGGCGCGGCGAGGAGAAACCATACAGACTTTTGTGACTGCGGTTGAGTACGTTGTGG 131  
|||||  
QY 133 AAGGAAAACTGTGCGCATTTCTAATTTGAAATCATCATCATCATCATCATCATCATCATCATCT 192  
|||||  
Db 132 AAGGAAAACTGTGCGCATTTCTAATTTGAAATCATCATCATCATCATCATCATCATCATCT 191  
|||||  
QY 193 GTTAACTGCTAACTTTTCTGTAACCAACCTGAGTCAACACAGATGAAATCCCTGTATTGAC 252  
|||||  
Db 192 GTTAACTGCTAACTTTTCTGTAACCAACCTGAGTCAACACAGATGAAATCCCTGTATTGAC 251  
|||||  
QY 253 ATTTAAAGATAATTTCTAAGTATGGTACCTTTGTTAATGAGGAAAAATGACAGAAATGGCTT 312  
|||||  
Db 252 ATTTAAAGATAATTTCTAAGTATGGTACCTTTGTTAATGAGGAAAAATGACAGAAATGGCTT 311  
|||||  
QY 313 TTCCCGAACTTTGAAGTCGGGGATGATATTAATTTTTCGAGTGTGTTGGAAGTAAATTCAG 372  
|||||  
Db 312 TTCCCGAACTTTGAAGTCGGGGATGATATTAATTTTTCGAGTGTGTTGGAAGTAAATTCAG 371  
|||||  
QY 373 AATAGAGTATGAGCCCTTTGGTTGCAATGCTCTTCTTTGTAGATGCTCTCTGGAAAACTGC 432  
|||||  
Db 372 AATAGAGTATGAGCCCTTTGGTTGCAATGCTCTTCTTTGTAGATGCTCTCTGGAAAACTGC 431  
|||||  
QY 433 TTTAAATCAAGCTATATTGCAACTTTGGAGATTTACTGTAAACAATTTGACAGAGAATG 492  
|||||  
Db 432 TTTAAATCAAGCTATATTGCAACTTTGGAGATTTACTGTAAACAATTTGACAGAGAATG 491  
|||||  
QY 493 CACTCACCTTGTTCATGGTATCAGTCAAAAGTTACCAATTAACCAATATATGTCACCTCATTTG 552  
|||||  
Db 492 CACTCACCTTGTTCATGGTATCAGTCAAAAGTTACCAATTAACCAATATATGTCACCTCATTTG 551  
|||||  
QY 553 TGGAGCTCCAAATTTGTAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA 612  
|||||  
Db 552 TGGAGCTCCAAATTTGTAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA 611  
|||||  
QY 613 GAAGCAGCTCCAAATTTGTAAGCTTTTACCACCTCTTGTATGAACCATCTATTGGAG 672  
|||||  
Db 612 GAAGCAGCTCCAAATTTGTAAGCTTTTACCACCTCTTGTATGAACCATCTATTGGAG 671  
|||||  
QY 673 TAAAAATGTTGATCTGTGAGGACGCGAGGAGAAAAAATAATCTTTCAAGGAAAAACATT 732  
|||||  
Db 672 TAAAAATGTTGATCTGTGAGGACGCGAGGAGAAAAAATAATCTTTCAAGGAAAAACATT 731  
|||||  
QY 733 TATATTTTGAATGCCAAACACGATAAGAAATTTGAGTTCCCGAGTTGCTTTTGGAGGTGG 792  
|||||

Db 732 TATATTTTGAATGCCAAA CAGCNAAGAAATTTGAGTTCGCGAGTGTCTTTGGAGGTGG 791  
Qy 793 GGAAGCTAGGTTGATTAACAGAGAGAGATGAAGAGAA CATAAATTTCTTTTGGCTCCGGG 852  
Db 792 GGAAGCTAGGTTGATTAACAGAGAGAGATGAAGAGAA CATAAATTTCTTTTGGCTCCGGG 851  
Qy 853 AAGCTGTGTTGATTAACAGAGATGAAGAGAA CATAAATTTCTTTTGGCTCCGGG 912  
Db 852 AAGCTGTGTTGATTAACAGAGATGAAGAGAA CATAAATTTCTTTTGGCTCCGGG 911  
Qy 913 GAAATGATTCAGTCAATATGATATGCTTCCAAAGGCAAGGCTTTAGACCTATTTCTGA 972  
Db 912 GAAATGATTCAGTCAATATGATATGCTTCCAAAGGCAAGGCTTTAGACCTATTTCTGA 971  
Qy 973 AGCAGAAATGGATTTGGCGGTGATTTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG 1032  
Db 972 AGCAGAAATGGATTTGGCGGTGATTTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG 1031  
Qy 1033 CCATCCAGTACAGGATTTAAAGACAACACTCCAGAGCCAAAGCTTTTCAAGGCGGTGC 1092  
Db 1032 CCATCCAGTACAGGATTTAAAGACAACACTCCAGAGCCAAAGCTTTTCAAGGCGGTGC 1091  
Qy 1093 AGTTGATGAAGAACTAATGCTCAAGCGCCCGAGTGAACTACAACTACAGTGTGACAC 1152  
Db 1092 AGTTGATGAAGAACTAATGCTCAAGCGCCCGAGTGAACTACAACTACAGTGTGACAC 1151  
Qy 1153 AGAATCAGAGCAAGCAGATACATGGGATTTGAGTGAAAGGCCAAAGAAATCAAAGTCTC 1212  
Db 1152 AGAATCAGAGCAAGCAGATACATGGGATTTGAGTGAAAGGCCAAAGAAATCAAAGTCTC 1211  
Qy 1213 CAAATGGAAACAAAATTCAGATGCTTTTCAAGAGCGACCCACTGTAAAGAGTCTTG 1272  
Db 1212 CAAATGGAAACAAAATTCAGATGCTTTTCAAGAGCGACCCACTGTAAAGAGTCTTG 1271  
Qy 1273 CAAACAGGCTCTAATAATAGTATGATGATCAATACCTTTGGCTAAGATGAGATCCC 1332  
Db 1272 CAAACAGGCTCTAATAATAGTATGATGATCAATACCTTTGGCTAAGATGAGATCCC 1331  
Qy 1333 AAATCTATCAGCTTTCCAACTAAATTTGCCAAGTATATAATAAAGTAAAGATAGGCTTC 1392  
Db 1332 AAATCTATCAGCTTTCCAACTAAATTTGCCAAGTATATAATAAAGTAAAGATAGGCTTC 1391  
Qy 1393 TCAGCAGCAGCAGCAACCTCCATCAGAACTATCTTTAGCCGCTCTACCAAAAAAGGGA 1452  
Db 1392 TCAGCAGCAGCAGCAACCTCCATCAGAACTATCTTTAGCCGCTCTACCAAAAAAGGGA 1451  
Qy 1453 AAGGATGAAGAAATCAAGAAATGCTTCATGCAATCAGCAAGATGAAACGCTTG 1512  
Db 1452 AAGGATGAAGAAATCAAGAAATGCTTCATGCAATCAGCAAGATGAAACGCTTG 1511  
Qy 1513 TTCTCTTTTGAACAAACACAACTGCTACACCTCATTTGTGGAAAAATTAAGGAGCAGCA 1572  
Db 1512 TTCTCTTTTGAACAAACACAACTGCTACACCTCATTTGTGGAAAAATTAAGGAGCAGCA 1571  
Qy 1573 TCTATCTGAGATGAGCCTGTGACACAACTCAGCAATTAATTTTACAGATACAGA 1632  
Db 1572 TCTATCTGAGATGAGCCTGTGACACAACTCAGCAATTAATTTTACAGATACAGA 1631  
Qy 1633 TTTTAAATCTATTTGAAAAATTTGCCAGTAATCTCATGCTGCAGAAAGCTAGATC 1692  
Db 1632 TTTTAAATCTATTTGAAAAATTTGCCAGTAATCTCATGCTGCAGAAAGCTAGATC 1691  
Qy 1693 AAATAAANAAGGAAATGGATGATGTGGCCATAGAAAGATGAAGTATTTGAAACAGTTATT 1752  
Db 1692 AAATAAANAAGGAAATGGATGATGTGGCCATAGAAAGATGAAGTATTTGAAACAGTTATT 1751  
Qy 1753 CAAGGACACAAAACAGAGTTAGAAATTTGATGTGAAGTTCAAAGGAGGAGATGT 1812  
Db 1752 CAAGGACACAAAACAGAGTTAGAAATTTGATGTGAAGTTCAAAGGAGGAGATGT 1811  
Qy 1813 CAATGTTAGAAAAGGCCAAGGATGGATATAGAAA CAAATGACACTTTTCAAGTGAAGC 1872  
Db 1812 CAATGTTAGAAAAGGCCAAGGATGGATATAGAAA CAAATGACACTTTTCAAGTGAAGC 1871

Qy 1873 AGTACCAGAAAGTAGCAAAATATCTCAAGAAATGAAATTTGGAGAAAAACGTGAACCTCAA 1932  
Db 1872 AGTACCAGAAAGTAGCAAAATATCTCAAGAAATGAAATTTGGAGAAAAACGTGAACCTCAA 1931  
Qy 1933 GGAAGACTCACTATGCTAGCTTAAGAAATATCTAACAATGACAAACCTTCAGGATGATAG 1992  
Db 1932 GGAAGACTCACTATGCTAGCTTAAGAAATATCTAACAATGACAAACCTTCAGGATGATAG 1991  
Qy 1993 TGAGATGCTTCCAAAAAAGCTGTTATTGACTGAATTTAGATCACTGCTGATTAATAAACCTC 2052  
Db 1992 TGAGATGCTTCCAAAAAAGCTGTTATTGACTGAATTTAGATCACTGCTGATTAATAAACCTC 2051  
Qy 2053 TACTTCCAGAAATCCGCTCTGGCATAAATGATGATTAATGCTCAACTAAAAAATTTCAAGAA 2112  
Db 2052 TACTTCCAGAAATCCATCTGGCATAAATGATGATTAATGCTCAACTAAAAAATTTCAAGAA 2111  
Qy 2113 ATTCAAAAAGGTCACATATCTGGAGCAGGAAAACTTCCACACATCATTTGGAGGATCAGA 2172  
Db 2112 ATTCAAAAAGGTCACATATCTGGAGCAGGAAAACTTCCACACATCATTTGGAGGATCAGA 2171  
Qy 2173 TCTAATAGCTCATCTGCTCGAAAGAAATACAGAACTAGAAAGAGTGGCTAAGGCGAGAAAT 2232  
Db 2172 TCTAATAGCTCATCTGCTCGAAAGAAATACAGAACTAGAAAGAGTGGCTAAGGCGAGAAAT 2231  
Qy 2233 GGAGGTACAAAATCAAATCTGCAAAAAGAGAGTCTCTTCTGCTGATGATCTTTTATAGATACAA 2292  
Db 2232 GGAGGTACAAAATCAAATCTGCAAAAAGAGAGTCTCTTCTGCTGATGATCTTTTATAGATACAA 2291  
Qy 2293 TCTTATTTTAAAAAGGAGAGATTAATCTGAGGATTTTAAAAAGAGCCATGGAACAACTTC 2352  
Db 2292 TCTTATTTTAAAAAGGAGAGATTAATCTGAGGATTTTAAAAAGAGCCATGGAACAACTTC 2351  
Qy 2353 CTAGTAAGCATCTACTTTCAGGCCAACAGGTTATATGAATATATAGTGTATAGAACGAT 2412  
Db 2352 CTAGTAAGCATCTACTTTCAGGCCAACAGGTTATATGAATATATAGTGTATAGAACGAT 2411  
Qy 2413 TTAAGTTACAAATGTTTATGGCTAAATTTTAAATAAATAATGCAACAAACCTTTGATTTCT 2472  
Db 2412 TTAAGTTACAAATGTTTATGGCTAAATTTTAAATAAATAATGCAACAAACCTTTGATTTCT 2471  
Qy 2473 TTTGATATGAAACAAATGTTTGTCTGTTTTCAGGCTTTGCTCATTTGATCTTTTTCATTT 2532  
Db 2472 TTTGATATGAAACAAATGTTTGTCTGTTTTCAGGCTTTGCTCATTTGATCTTTTTCATTT 2531  
Qy 2533 TTTAAATGTTTGTGTTTATTAATAATAGTTAAATATAGTCAAGTTCAAAAATTTCTAAATATTA 2592  
Db 2532 TTTAAATGTTTGTGTTTATTAATAATAGTTAAATATAGTCAAGTTCAAAAATTTCTAAATATTA 2591  
Qy 2593 CGTAAGGTAAAGGACTAAAGTCAACCTTCCACCATTTGCTCTAGTACTTCTAGTACTTCT 2640  
Db 2592 CGTAAGGTAAA - GACTAAAGTCAACCTTCCACCATTTGCTCTAGTACTTCTAGTACTTCT 2650  
Qy 2641 ----- 2640  
Db 2651 GAAAAAATTCATGATCTATCTTCTTATGAATCTTTCCAGGATTTTTCAGTCTCTATTTCT 2710  
Qy 2641 ----- TATTTTAAATAATTTTCTACAAATGATAGCAATAACATATGCAAGTGTCTA 2693  
Db 2711 AAATTCCTATTTTAAATTAATTTTCTACAAATGATAGCAATAACATATGCAAGTGTCTA 2770  
Qy 2694 CACCTTGTCTTTTACTTAGTAAGATTTAAAAATTTATAGGAATATCAATAATATGTTTTTTA 2753  
Db 2771 CACCTTGTCTTTTACTTAGT - AGATTTAAAAATTTATAGGAATATCAATAATATGTTTTTTA 2829  
Qy 2754 ATATTTTCTTTTCCATTTGCTGTAGTCTTACCTAAACTCTGGTGATCCAAAACAAAT 2813  
Db 2830 ATATTTTCTTTTCCATTTGCTGTAGTCTTACCTAAACTCTGGTGATCCAAAACAAAT 2889  
Qy 2814 GGCCTCAGTGCGAGATGTCACCTACATGTTATTTCTAGTACTAGAACTGAGAACCATG 2873  
Db 2890 GGCCTCAGTGCGAGATGTCACCTACATGTTATTTCTAGTACTAGAACTGAGAACCATG 2949



QY	253	ATTAAAGATAAATCTTAAGTATGGTACCTTTGTTAATAGAGAAATAATGCGAATATGCGTT	312
DB	241	ATTAAAGATAAATCTTAAGTATGGTACCTTTGTTAATAGAGAAATAATGCGAATATGCGTT	300
QY	313	TTCCCGAATCTTGAAGTCGGGGATGGTATTACTTTTGGAGTGTTTGGAGTAAATTCAG	372
DB	301	TTCCCGAATCTTGAAGTCGGGGATGGTATTACTTTTGGAGTGTTTGGAGTAAATTCAG	360
QY	373	AATAGAGTATGAGCCCTTTGGTTGCATGCTCTTCTTGTTTAGATGCTCTCGGAAAAATGC	432
DB	361	AATAGAGTATGAGCCCTTTGGTTGCATGCTCTTCTTGTTTAGATGCTCTCGGAAAAATGC	420
QY	433	TTTAAATCAAGTATATTGCAACTTGGAGGATTTACTGTAAACAATTTGAGCAGAAGAATG	492
DB	421	TTTAAATCAAGTATATTGCAACTTGGAGGATTTACTGTAAACAATTTGAGCAGAAGAATG	480
QY	493	CACTCACCTCTGTCATGGTATCAGTGAAGTTTACCATTAAAAACAATATGTGCACCTCATTTG	552
DB	481	CACTCACCTCTGTCATGGTATCAGTGAAGTTTACCATTAAAAACAATATGTGCACCTCATTTG	540
QY	553	TGGACGTCCAATTGTAAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA	612
DB	541	TGGACGTCCAATTGTAAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA	600
QY	613	GAAGCAGCCTCCACAATAATTGAAAGTTTTTACCACCTCTTGATGAAACAATCTATTGGAAG	672
DB	601	GAAGCAGCCTCCACAATAATTGAAAGTTTTTACCACCTCTTGATGAAACAATCTATTGGAAG	660
QY	673	TAAAAATGTTGATCTGTCAGACGGCAGGAAGAAACAAATCTTCAAGGGAAAAACATT	732
DB	661	TAAAAATGTTGATCTGTCAGACGGCAGGAAGAAACAAATCTTCAAGGGAAAAACATT	720
QY	733	TATATTTTGAATGCCAAACAGCATAAGAAATTCAGTTCCGCAGTTGTCTTTGGAGGTGG	792
DB	721	TATATTTTGAATGCCAAACAGCATAAGAAATTCAGTTCCGCAGTTGTCTTTGGAGGTGG	780
QY	793	GGAACTAGTGTGATTAACAGAAAGAGAATGAAGAAAGAAACATAATTTCTTTTGGCTCCGGG	852
DB	781	GGAACTAGTGTGATTAACAGAAAGAGAATGAAGAAAGAAACATAATTTCTTTTGGCTCCGGG	840
QY	853	AACGTGTGTTGATACAGAAATAACAACTCAGACCTTAAATTCCTGACTGTGCAGAA	912
DB	841	AACGTGTGTTGATACAGAAATAACAACTCAGACCTTAAATTCCTGACTGTGCAGAA	900
QY	913	GAATGGATTCAATTAATGGATATGCTCCAAAGGCAAGGCTCTTAGACCTATTCTCTGA	972
DB	901	GAATGGATTCAATTAATGGATATGCTCCAAAGGCAAGGCTCTTAGACCTATTCTCTGA	960
QY	973	AGCAGAAATCGAATGGCGGTGATTTTTCATGACTACAAAGAAATTAATCTGTATCCTCAGGG	1032
DB	961	AGCAGAAATCGAATGGCGGTGATTTTTCATGACTACAAAGAAATTAATCTGTATCCTCAGGG	1020
QY	1033	CCATCCCAGTACAGATTAAAGACAAACAACTCAGGACCAAGCCCTTTTCAACAGCGTGTCT	1092
DB	1021	CCATCCCAGTACAGATTAAAGACAAACAACTCAGGACCAAGCCCTTTTCAACAGCGTGTCT	1080
QY	1093	AGTTGATGAAAACTAATGCCCAGCGCCCAAGTGAAACAACACTACAACTACGTAGCTGCACAC	1152
DB	1081	AGTTGATGAAAACTAATGCCCAGCGCCCAAGTGAAACAACACTACAACTACGTAGCTGCACAC	1140
QY	1153	AGAAATCAGACGACGATACATCGGATTTGAGTGAAAGGCCAAAGAAATCAAAGTCTC	1212
DB	1141	AGAAATCAGACGACGATACATCGGATTTGAGTGAAAGGCCAAAGAAATCAAAGTCTC	1200
QY	1213	CAAAATGGAAACAAAATTCAGAAATGCTTTTCAACAGACGCCACCTGTAAGGAGTCTTG	1272
DB	1201	CAAAATGGAAACAAAATTCAGAAATGCTTTTCAACAGACGCCACCTGTAAGGAGTCTTG	1260
QY	1273	CAAAACAGCTCTAATAATAATAGTATGGTATCAAAATCTTTGCTTAAGATGAGAATCCC	1332
DB	1261	CAAAACAGCTCTAATAATAATAGTATGGTATCAAAATCTTTGCTTAAGATGAGAATCCC	1320
QY	1333	AAACTATCAGCTTTTCCACCACTAAATTTGCCAAGTATAAATAAAAGTAAAGATAGGGCTTC	1392

[illegible]

Db 2401 TTAAGTTACAATGTTTTATGGCCTAAATTTATTAATAAAATGACAAAACTTTGATCT 2460  
Qy 2473 TTTGTATGTAACAAATTTGTTGTCGTTTTTCAGGCTTTGTCATGTCATCTTTTTCATT 2532  
Db 2461 TTTGTATGTAACAAATTTGTTGTCGTTTTTCAGGCTTTGTCATGTCATCTTTTTCATT 2520  
Qy 2533 TTTAAATGTTGTTGTTTAAATAGTTAAATAGTTAAATAGTCACAGTTCAAAAATCTTAAATRTA 2592  
Db 2521 TTTAAATGTTGTTGTTTAAATAGTTAAATAGTTAAATAGTCACAGTTCAAAAATCTTAAATATA 2580  
Qy 2593 CGTAAGGTAAGGACTAAAGTCAACCTTCCACCAATGTCCTAGCTACT 2640  
Db 2581 CGTAAGGTAAGGACTAAAGTCAACCTTCCACCAATGTCCTAGCTACT 2639  
Qy 2641 ----- 2640  
Db 2640 GAAAAAAATTCATGACTACTCTTATGAAATCTTTCCAGGGATTTTTGAGTCCCTATTTC 2699  
Qy 2641 -----TATTTTTAAATAATTTCCCTACACAAATGATAGATAACATATGCAAGTGTCTA 2693  
Db 2700 AAATTCCTATTTTAAATAATTTCCCTACACAAATGATAGATAACATATGCAAGTGTCTA 2759  
Qy 2694 CACCTTGCTTTTACTTAGTAAGATTAAGAAATTAATAGGAATATCAATAATATGTTTTTA 2753  
Db 2760 CACCTTGCTTTTACTTAGT-AGATTAAGAAATTAAGGAATATCAATAATATGTTTTTA 2818  
Qy 2754 ATATTTTTTCTTTTCCATATGCTGTAGTCTTACCTAAACTCTGGTGATCCAAACAAAT 2813  
Db 2819 ATATTTTTTCTTTTCCATATGCTGTAGTCTTACCTAAACTCTGGTGATCCAAACAAAT 2878  
Qy 2814 GGCCTCAGTGGTGACAGATGTCACCTACATGTTATCTAGTACTAGAAATCGAAGACCATG 2873  
Db 2879 GGCCTCAGTGGTGACAGATGTCACCTACATGTTATCTAGTACTAGAAATCGAAGACCATG 2938  
Qy 2874 TGGAGACTTCATCAAAACATGGGTTAGTTTTCACCAAGATGGAAGACCTGTACCCCTTT 2933  
Db 2939 TGGAGACTTCATCAAAACATGGGTTAGTTTTCACCAAGATGGAAGACCTGTACCCCTTT 2998  
Qy 2934 TTGGTGGTCTTACTAGCTGGGTTGGTGCTGTTTTGAGCTTATTTAGAGTCTTAGTCTTT 2993  
Db 2999 TTGGTGGTCTTACTAGCTGGGTTGGTGCTGTTTTGAGCTTATTTAGAGTCTTAGTCTTT 3058  
Qy 2994 CCTACTTATAAGTAGAATGTTGAGATGTTTTCTTTTCTTACCKTAAAGGGAGATGTT 3053  
Db 3059 CCTACTTATAAGTAGAATGTTGAGATGTTTTCTTTCTTACCTTAAAGGGAGATGTT 3118  
Qy 3054 AAGAAAACATGAATGCTTTTTTCAAACCTTATATGACAGTGAATTTTCAAGTCTGTGTTTC 3113  
Db 3119 AAGAAAACATGAATGCTTTTTTCAAACCTTATATGACAGTGAATTTTCAAGTCTGTGTTTC 3178  
Qy 3114 AAAAAATATTCATGTACCTGTGATCCAGCAAGAGGGAGTTCCAGTCAAGATCACTATC 3173  
Db 3179 AAAAAATATATTCATGTACCTGTGATCCAGCAAGAGGGAGTTCCAGTCAAGATCACTATC 3238  
Qy 3174 AACTGATTAAGTCTTTTAGAGAATGAGAAATGGAACAGTGGAGGAATCGAGGCCATATTTCC 3233  
Db 3239 AACTGATTAAGTCTTTTAGAGAATGAGAAATGGAACAGTGGAGGAATCGAGGCCATATTTCC 3298  
Qy 3234 ATGACTTCCCTGTAAACAGAACCAACAGAAAGGCAAGAGGCTGGCCCTCATACACTC 3293  
Db 3299 ATGACTTCCCTGTAAACAGAACCAACAGAAAGGCAAGAGGCTGGCCCTCATACACTC 3358  
Qy 3294 TCACCTTCCAAATCTTTGTGGAGTGCATCTACTTGCAGAACCAAAATTAACCTTACTTCCA 3353  
Db 3359 TCACCTTCCAAATCTTTGTGGAGTGCATCTACTTGCAGAACCAAAATTAACCTTACTTCCA 3418  
Qy 3354 AGTTCTCGGCTGCTTGACAGTGGAACTCCAGCTCAGAGGGAGTTAGGGAATGAAGTCTTT 3413  
Db 3419 AGTTCTCGGCTGCTTGACAGTGGAACTCCAGCTCAGAGGGAGTTAGGGAATGAAGTCTTT 3478  
Qy 3414 TTTTAAAAAGCTTCTCAGCCCTTCTAGGGGAACAGAAATTTGGGTGAGCCAAATCTGCAATTT 3473  
Db 3479 TTTTAAAAAGCTTCTCAGCCCTTCTAGGGGAACAGAAATTTGGGTGAGCCAAATCTGCAATTT 3538

Qy 3474 CTACTACAGGCATTCGAGACCAGTTAGATTATTGAAATATTATAGAGAGTTATGAACACTT 3533  
Db 3539 CTACTACAGGCATTCGAGACCAGTTAGATTATTGAAATATTATAGAGAGTTATGAACACTT 3598  
Qy 3534 AAATTTATGATAGTGGTATGACATTCGGATAGAACATGGGATACCTTTAGAGTAGAATTGAC 3593  
Db 3599 AAATTTATGATAGTGGTATGACATTCGGATAGAACATGGGATACCTTTAGAGTAGAATTGAC 3658  
Qy 3594 AGGCATATTTAGTTGATGAAATGGAGTCATTTGAGTCTTTTAAATAGCCATGTATCATAT 3653  
Db 3659 AGGCATATTTAGTTGATGAAATGGAGTCATTTGAGTCTCTTAAATAGCCATGTATCATAT 3718  
Qy 3654 TACCAGGTGAAGCTGCTGGAACATATGCTCCATTTTACAGTTTAAAGAAATATAATGAC 3713  
Db 3719 TACCAGGTGAAGCTGCTGGAACATATGCTCCATTTTACAGTTTAAAGAAATATAATGAC 3778  
Qy 3714 AGATTAAATATTGTTTCTGTCATGCCCAATCCCTTTCTAAGGAAGACTGCCCCTACTAT 3773  
Db 3779 AGATTAAATATTGTTTCTGTCATGCCCAATCCCTTTCTAAGGAAGACTGCCCCTACTAT 3838  
Qy 3774 AGCAGTTTTTATATTTGTCATTTTAAATGAATATAATGAATGAGGAGTCTGGTACCTCCTG 3833  
Db 3839 AGCAGTTTTTATATTTGTCATTTTAAATGAATATAATGAATGA-GAGTCTGGTACCTCCTG 3897  
Qy 3834 TCTTTTACAAATATTGCGGTGTTGTCAGTATTTTCCCTTTTAAACCMWTTCCCAATTCGGG 3893  
Db 3898 TCTTTTACAAATATTGCGGTGTTGTCAGTATTTTCCCTTTTAAACCMWTTCCCAATTCGGG 3953  
Qy 3894 TGTGTAGGTGAGTGTTCCTTTGCGGTGTTTAAATTTGATATATCCCTGATAGCTATTAATGG 3953  
Db 3954 TGTGTAGGTGAGTGTTCCTTTGCGGTGTTTAAATTTGATATATCCCTGATAGCTATTAATGG 4008  
Qy 3954 GTCATAGAAATCTTTTATACATTTCTAGATGCAAGTCTCTGTCGATATACGATATTGAGA 4013  
Db 4009 GTCATAGAAATCTTTTATACATTTCTAGATGCAAGTCTCTGTCGATATACGATATTGAGA 4068  
Qy 4014 TATTACACTAGTCTGTCGCTTGACTGTTTTCTTTATGTCCTTTTATGATGAATAGAAAGTTTT 4073  
Db 4069 TATTACACTAGTCTGTCGCTTGACTGTTTTCTTTATGTCCTTTTATGATGAATAGAAAGTTTT 4128  
Qy 4074 AAATTTTGAAGGTCACAAATTTATTTTCTTTTGTGATATTTTCTCTCCAAATTT 4133  
Db 4129 AAATTTTGAAGGTCACAAATTTATTTTCTTTTGTGATATTTTCTCTCCAAATTT 4188  
Qy 4134 RACCCCAAGNTTCAGATATTTCTGCTCTATATATATAAACTTTATATTTTATATTGAGA 4193  
Db 4189 RACCCCAAGNTTCAGATATTTCTGCTCTATATATATAAACTTTATATTTTATATTGAGA 4248  
Qy 4194 TCTACCTTGAATTTGATATGTTGTGAATTTATGATCAGGTTCTTTTTCCTCCCAT 4253  
Db 4249 TCTACCTTGAATTTGATATGTTGTGAATTTATGATCAGGTTCTTTTTCCTCCCAT 4308  
Qy 4254 ACAAGTATCCAGTCAATGTAACACTGTTTATTTGAAAGAAATTAATCTTCTCATTAATTT 4313  
Db 4309 ACAAGTATCCAGTCAATGTAACACTGTTTATTTGAAAGAAATTAATCTTCTCATTAATTT 4368  
Qy 4314 ACCTTGGCAATTAGTAAAAA 4333  
Db 4369 ACCTTGGCAATTAGTAAAAA 4388

## RESULT 4

AR233599

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AR233599 Sequence 1 from patent US 6458534.

AR233599

AR233599.1 GI:2726201

Unknown.

Unclassified.

1 (bases 1 to 4386)

4386 bp

DNA

linear

PAT 20-DEC-2002

AUTHORS Concannon, P.J., Vissinga, C.S., Cerosaletti, K.M., Varon-Mateeva, R.,  
Sperling, K. and Reis, A.W.S.  
TITLE Gene associated with NiJmegen breakage syndrome, it's gene product  
and methods for their use  
JOURNAL Patent: US 6458534-A 1 01-OCT-2002;  
FEATURES Location/Qualifiers  
source i. 4396  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 93.38; Score 4108.4; DB 6; Length 4386;  
Best Local Similarity 97.19; Pred. No. 0;  
Matches 4272; Conservative 7; Mismatches 28; Indels 91; Gaps 6;  
QY 28 GCCCCAGCCCTGAGGAGCCGCGATGTGGAATCTGCTCCCGCCCGCGCGCGCGGCGG 87  
DB 1 GCCCCAGCCCTGAGGAGCCGCGATGTGGAATCTGCTCCCGCCCGCGCGCGCGGCGG 60  
QY 88 AGGAGAACCATACAGACTTTTGACTGGCGTTGAGTACGTTGTTGGAAGGAAAACTGTGC 147  
DB 61 AGGAGAACCATACAGACTTTTGACTGGCGTTGAGTACGTTGTTGGAAGGAAAACTGTGC 120  
QY 148 CATTCTTAATTGAAATGATCAGTGCATCAGCCGGAATCATGCTGTGTTAACTGCTAACTT 207  
DB 121 CATTCTGATTGAAATGATCAGTGCATCAGCCGGAATCATGCTGTGTTAACTGCTAACTT 180  
QY 208 TTCTGTAAACCACTGAGTCAAAACAGATGAAATCCCTGTATGACATTTAAAGATAATTC 267  
DB 181 TTCTGTAAACCACTGAGTCAAAACAGATGAAATCCCTGTATGACATTTAAAGATAATTC 240  
QY 268 TAAGTATGTGACCTTTGTTAATGAGGAAAAAATGCAAGTGGCTTTTCCGAACTTTGAA 327  
DB 241 TAAGTATGTGACCTTTGTTAATGAGGAAAAAATGCAAGTGGCTTTTCCGAACTTTGAA 300  
QY 328 GTCCGGGGATGTTTACTTTTGGAGTGTGTTGGAAGTAAATTCAGATAGATGAGCC 387  
DB 301 GTCCGGGGATGTTTACTTTTGGAGTGTGTTGGAAGTAAATTCAGATAGATGAGCC 360  
QY 388 TTTGGTTGCAATGCTCTCTGTTTGTAGATGCTCTGGGAAATGCTTTAATCAAGCTAT 447  
DB 361 TTTGGTTGCAATGCTCTCTGTTTGTAGATGCTCTGGGAAATGCTTTAATCAAGCTAT 420  
QY 448 ATTGCAACTTTGGAGGATTTACTGTAAACAAATTGGACAGAAATGCACTCACCTTTGTCAT 507  
DB 421 ATTGCAACTTTGGAGGATTTACTGTAAACAAATTGGACAGAAATGCACTCACCTTTGTCAT 480  
QY 508 GGTATCAGTGAAAGTTACCAATTAACAAATATGTGCACTCATTTTGGAGCTCCAATGTT 567  
DB 481 GGTATCAGTGAAAGTTACCAATTAACAAATATGTGCACTCATTTTGGAGCTCCAATGTT 540  
QY 568 AAAGCCAGAAATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCCAAGAGCAGCTCCACA 627  
DB 541 AAAGCCAGAAATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCCAAGAGCAGCTCCACA 600  
QY 628 AATTGAAAGTTTAAACCCACTCTTGAATGAACCATCTATTGGAAGTAAATAATGTTGATCT 687  
DB 601 AATTGAAAGTTTAAACCCACTCTTGAATGAACCATCTATTGGAAGTAAATAATGTTGATCT 660  
QY 688 GTACAGCGCAGGAAAGAAACAAATCTTCAAGGGAACAAATTTATTTTGAATGC 747  
DB 661 GTACAGCGCAGGAAAGAAACAAATCTTCAAGGGAACAAATTTATTTTGAATGC 720  
QY 748 CAAACAGCATTAAGAAATTTGAGTTCCGAGTTGCTTTGGAGGTGGGAAAGCTAGGTTGAT 807  
DB 721 CAAACAGCATTAAGAAATTTGAGTTCCGAGTTGCTTTGGAGGTGGGAAAGCTAGGTTGAT 780  
QY 808 AACAGAGGAATGAAGAGAAACATAATTTCTTTTGGCTCCGGGAACGTTGTTGTTGA 867  
DB 781 AACAGAGGAATGAAGAGAAACATAATTTCTTTTGGCTCCGGGAACGTTGTTGTTGA 840  
QY 868 TACAGGAATAAACAACTCAGACCTTTAATCTCTGATCTGTTCAGAGAAATGGATTCAGTC 927

DB 841 TACAGGAATAAACAACTCAGACACCTTAATTCCTGACTGTTCAGAGAAATGGATTCAGTC 900  
QY 928 AATAATGGATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTGTAAGACAGAAATGGATTC 987  
DB 901 AATAATGGATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTGTAAGACAGAAATGGATTC 960  
QY 988 GCGGGTGAATTTTCATGACTACAAAGAAATTACTGTGATTCCTCAGGGGCATCCAGTACAGG 1047  
DB 961 GCGGGTGAATTTTCATGACTACAAAGAAATTACTGTGATTCCTCAGGGGCATCCAGTACAGG 1020  
QY 1048 ATTAAGACAAACAACTCCAGGACCAAGCCTTTTCAAGGCGGTGTGAGTTGATCAAAAATCT 1107  
DB 1021 ATTAAGACAAACAACTCCAGGACCAAGCCTTTTCAAGGCGGTGTGAGTTGATCAAAAATCT 1080  
QY 1108 AATGCCAAGCGCCCGCCAGTGAACACATACATACATACATACATACATACATACATACAT 1167  
DB 1081 AATGCCAAGCGCCCGCCAGTGAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 1140  
QY 1168 AGATACATGGGATTTGAGTGAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 1227  
DB 1141 AGATACATGGGATTTGAGTGAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCA 1200  
QY 1228 ATTCAGAAATGCTTTTCAAGAGCGCACCTCTGTAAAGGAGTCTCTGCAAAACAGCTCTAA 1287  
DB 1201 ATTCAGAAATGCTTTTCAAGAGTGCACCCACTGTAAAGGAGTCTCTGCAAAACAGCTCTAA 1260  
QY 1288 TAATAATAGTATGGTATCAATACATTTGGCTAAGATGAGAAATCCCAAACTATCAGGCTTTC 1347  
DB 1261 TAATAATAGTATGGTATCAATACATTTGGCTAAGATGAGAAATCCCAAACTATCAGGCTTTC 1320  
QY 1348 ACCAACTAAATGCGCAAGTATATAATAAAGTAAAGATAGGGCTTTCTCAGCAGCAGCAGAC 1407  
DB 1321 ACCAACTAAATGCGCAAGTATATAATAAAGTAAAGATAGGGCTTTCTCAGCAGCAGCAGAC 1380  
QY 1408 CAACTCCATCAGAACTACTTTCCGCGCTCTACCAAAAAAGGAAAGGAAAGGAAAGGAA 1467  
DB 1381 CAACTCCATCAGAACTACTTTCCGCGCTCTACCAAAAAAGGAAAGGAAAGGAAAGGAA 1440  
QY 1468 TCAAGAAATGCTTTTCATGCAAAATCAGCAAGATAGAAACGCTTTGTTCTCTTTTGAACA 1527  
DB 1441 TCAAGAAATGCTTTTCATGCAAAATCAGCAAGATAGAAACGCTTTGTTCTCTTTTGAACA 1500  
QY 1528 AACACAACTGCTTACACCTCATTTGFGAAAAAATAGGAGCAGCATCTATCTGAGAAATGA 1587  
DB 1501 AACACAACTGCTTACACCTCATTTGFGAAAAAATAGGAGCAGCATCTATCTGAGAAATGA 1560  
QY 1588 GCTGTGGACACAACTCAGCAATTAATCTTTTACAGATACAGATTTAAATCTATTCT 1647  
DB 1561 GCTGTGGACACAACTCAGCAATTAATCTTTTACAGATACAGATTTAAATCTATTCT 1620  
QY 1648 GAAAAATTCGCCAGTAAATCTCATGCTGCAGAAAGCTTAAGATCAAAATAAAAAAGGGA 1707  
DB 1621 GAAAAATTCGCCAGTAAATCTCATGCTGCAGAAAGCTTAAGATCAAAATAAAAAAGGGA 1680  
QY 1708 AATGATGATGTGGCCATAGAAATGAAGTATTTGAAACAGTTATTTCAGGACACAAAAAC 1767  
DB 1681 AATGATGATGTGGCCATAGAAATGAAGTATTTGAAACAGTTATTTCAGGACACAAAAAC 1740  
QY 1768 AGAGTTAGAAATTTGATGCAAAAGTTTCAAAAACAGGAGGAGATGTCAATGTTAGAAAAAG 1827  
DB 1741 AGAGTTAGAAATTTGATGCAAAAGTTTCAAAAACAGGAGGAGATGTCAATGTTAGAAAAAG 1800  
QY 1828 GCCAAGGATGGATATAGAAAACAAATGACACTTTTTCAGTGTATGAGCAGTACCAGAAAGTAG 1887  
DB 1801 GCCAAGGATGGATATAGAAAACAAATGACACTTTTTCAGTGTATGAGCAGTACCAGAAAGTAG 1860  
QY 1888 CAAAAATATCTCAAGAAAAATGAAATTTGGGAAGAAAACGTGAACCTCAAGGAAGACTCACTATG 1947  
DB 1861 CAAAAATATCTCAAGAAAAATGAAATTTGGGAAGAAAACGTGAACCTCAAGGAAGACTCACTATG 1920  
QY 1948 GTCAAGTAAAGAAATATCTTAAACATGACAACTTCCAGGATGATAGTACAGATGCTTCCAAA 2007  
DB 1921 GTCAAGTAAAGAAATATCTTAAACATGACAACTTCCAGGATGATAGTACAGATGCTTCCAAA 1980

QY	2008	AAAGCTGTTATTGACCTGAATTTAGATCACTGGTGATTTAAAACTCTACTTCCAGAAATCC	2067	QY	3009	GAATGGTGAGATTGTTTTCTTTTCTACCKTAAAGGAGATGGTAAGAAAACAATGAATG	3068
DB	1981	AAAGCTGTTATTGACCTGAATTTAGATCACTGGTGATTTAAAACTCTACTTCCAGAAATCC	2040	DB	3059	GAATGGTGAGATTGTTTTCTTTTCTACCTTAAAGGAGATGGTAAGAAAACAATGAATG	3118
QY	2068	GTCGTCATAAAATGATGATTTATGGTCAACTTAAAAATTTCAAGAAATTTCAAAAAGGTCAC	2127	QY	3069	TCCTTTTTCAAACTTTATTGACAAGTGAATTTTCAAGTCTGTGTTTCAAAAATATATTCATG	3128
DB	2041	ATCTGGCATAAATGATGATTTATGGTCAACTTAAAAATTTCAAGAAATTTCAAAAAGGTCAC	2100	DB	3119	TCCTTTTTCAAACTTTATTGACAAGTGAATTTTCAAGTCTGTGTTTCAAAAATATATTCATG	3178
QY	2128	ATATCTGGAGCAGGAAAACTTCCACACATCATTTGAGGATCAGATCTAATAGCTCATCA	2187	QY	3129	TACCTGTGATCCAGCAAGAGGGAGTTCAGTCAAGAGTCACTACAACCTGATTGTT	3188
DB	2101	ATATCTGGAGCAGGAAAACTTCCACACATCATTTGAGGATCAGATCTAATAGCTCATCA	2160	DB	3179	TACCTGTGATCCAGCAAGAGGGAGTTCAGTCAAGAGTCACTACAACCTGATTGTT	3238
QY	2188	TGCTCCAAAGAAATACAGAACTAGAGAGTGGCTTAAGCGCAGGAAATGGAGTCAAAATCA	2247	QY	3189	TAGAAATGAGAAATGGAACAGTGAAGGATGAGGCCATATTTCCATGACTTCCCTTGTA	3248
DB	2161	TGCTCCAAAGAAATACAGAACTAGAGAGTGGCTTAAGCGCAGGAAATGGAGTCAAAATCA	2220	DB	3239	TAGAAATGAGAAATGGAACAGTGAAGGATGAGGCCATATTTCCATGACTTCCCTTGTA	3298
QY	2248	ACATGCAAAAGAAAGAGTCTCTGCTGATGATCTTTTATAGATAAATCCTTATTTAAAAAG	2307	QY	3249	AAACAGAAACACAGAAAGGACAAAGAGGCTGCTCTACATCACTCTCACCTTCCAAATCT	3308
DB	2221	ACATGCAAAAGAAAGAGTCTCTGCTGATGATCTTTTATAGATAAATCCTTATTTAAAAAG	2280	DB	3299	AAACAGAAACACAGAAAGGACAAAGAGGCTGCTCTACATCACTCTCACCTTCCAAATCT	3358
QY	2308	GAGAAATAACTGAGGATTTTAAAAAGAGCCATGGAAAAAATCTTCTAGTAAGCATCTAC	2367	QY	3309	TGTGAAAGTGCATCTACTTGCAGAACCAAAATTAATCTTACTTCCAAAGTCTGGCTGCTTG	3368
DB	2281	GAGAAATAACTGAGGATTTTAAAAAGAGCCATGGAAAAAATCTTCTAGTAAGCATCTAC	2340	DB	3359	TGTGAAAGTGCATCTACTTGCAGAACCAAAATTAATCTTACTTCCAAAGTCTGGCTGCTTG	3418
QY	2368	TTGAGGCCAACACAGGTTTATATGAATATATAGTCTATAGAGCGATTTAAGTTACATGTT	2427	QY	3369	CAGGTGGAACTCCAGCTCAAGGGAGTTAGGAAATGAAGTCTTTTTTAAAAAGCTCT	3428
DB	2341	TTGAGGCCAACACAGGTTTATATGAATATATAGTCTATAGAGCGATTTAAGTTACATGTT	2400	DB	3419	CAGGTGGAACTCCAGCTCAAGGGAGTTAGGAAATGAAGTCTTTTTTAAAAAGCTCT	3478
QY	2428	TTATGGCCATAAATTTTAAATAAAATGACAAAACTTTGATCTTTTGTATGTAAACAAT	2487	QY	3429	CAGCCTTCTAGGAAACAGAAATTTGGTGAGCAATCTGCAATTTCTACTACAGGCAATG	3488
DB	2401	TTATGGCCATAAATTTTAAATAAAATGACAAAACTTTGATCTTTTGTATGTAAACAAT	2460	DB	3479	CAGCCTTCTAGGAAACAGAAATTTGGTGAGCAATCTGCAATTTCTACTACAGGCAATG	3538
QY	2488	TGTTGTGTCGTTTTCAGGCTTTGTCATTGCAATCTTTTTCATTTTTTAAAAAGTGTGTTG	2547	QY	3489	AGACAGCTAGATTAATTAATTAATTAATAGAGATTTATGAACACTTAAATTTATGATAGT	3548
DB	2461	TGTTGTGTCGTTTTCAGGCTTTGTCATTGCAATCTTTTTCATTTTTTAAAAAGTGTGTTG	2520	DB	3539	AGACAGCTAGATTAATTAATTAATTAATAGAGATTTATGAACACTTAAATTTATGATAGT	3598
QY	2548	TTTATTAATAGTTAATATAGTACACAGTTTCAAAATTTCTAAATTTAGTAAAGTAAAGAC	2607	QY	3549	TATGACATTTGATAGAACATGGGATCTTTTAAAGTAGAATTTGACAGGCAATTTAGTTG	3608
DB	2521	TTTATTAATAGTTAATATAGTACACAGTTTCAAAATTTCTAAATTTAGTAAAGTAAAGAC	2579	DB	3599	TATGACATTTGATAGAACATGGGATCTTTTAAAGTAGAATTTGACAGGCAATTTAGTTG	3658
QY	2608	TAAAGTCAACCTTCCACCAATTTCTAGTACTT-----	2640	QY	3609	ATGAAATGGAGTCATTGAGTCTTTTAAATAGCCATGTATCATTAATTAACAAGTGAAGCTG	3668
DB	2580	TAAAGTCAACCTTCCACCAATTTCTAGTACTT-----	2639	DB	3659	ATGAAATGGAGTCATTGAGTCTTTTAAATAGCCATGTATCATTAATTAACAAGTGAAGCTG	3718
QY	2641	-----TATTTTAA	2648	QY	3669	GTGGAACATATGCTCCATTTTACAGTTTAAAGGAATATTAAGGACAGATTAATATGTTT	3728
DB	2640	TACTCATTTTCTTATGAATCTTTCCAGGGAATTTTGTAGTCCCTATTTCAAAATTTCTATTTTAA	2699	DB	3719	GTGGAACATATGCTCCATTTTACAGTTTAAAGGAATATTAAGGACAGATTAATATGTTT	3778
QY	2649	AATAATTTCTACACAAATGATAGCATACATATGAGTGTCTACACCTTGTCTTTTAA	2708	QY	3729	TCGTGATGCCCCACCAATCCCTTTCTTAAAGGAAGACTGCCCTACTATAGCAGTTTATATTT	3788
DB	2700	AATAATTTCTACACAAATGATAGCATACATATGAGTGTCTACACCTTGTCTTTTAA	2759	DB	3779	TCGTGATGCCCCACCAATCCCTTTCTTAAAGGAAGACTGCCCTACTATAGCAGTTTATATTT	3838
QY	2709	CTTAGTAAAGATTTAAATATATAGGAATATCAATATATGTTTAAATATTTTCTTTTC	2768	QY	3789	TGTCAATTTATGAATATTAATGAATGAGGATTTCTGGTACCTCTCTTTTACAAATATG	3848
DB	2760	CTTAGT-AGATTAATAATATATAGGAATATCAATATATGTTTAAATATTTTCTTTTC	2818	DB	3839	TGTCAATTTATGAATATTAATGAATGAGGATTTCTGGTACCTCTCTTTTACAAATATG	3897
QY	2769	CATTATGCTGATCTTACCTAACTCTGCTGATCCAAACAAATGGCTTCAGTGGTGA	2828	QY	3849	GCTGTTGCTCCAGTATTTTCCCTTTTAAACCMTTTCCCAATTCGGGTGTGAGTGGATCT	3908
DB	2819	CATTATGCTGATCTTACCTAACTCTGCTGATCCAAACAAATGGCTTCAGTGGTGA	2878	DB	3898	G-----TGTGTGCTGATTTTCTTTTAAACCMTTTCCCAATTCGGGTGTGAGTGGATCT	3948
QY	2829	GATGTCACCTACATGTTATCTAGTACTAGAACTGAAGCAAGCAATGTGGAGACTTCATCAA	2888	QY	3909	TTCCATTTTGGGTTTTAAATTTGTATATCCCTGATAGCTATAATTTGGGTCTATAGAAATTTCTT	3968
DB	2879	GATGTCACCTACATGTTATCTAGTACTAGAACTGAAGCAAGCAATGTGGAGACTTCATCAA	2938	DB	3949	TTTCAATTTGGTTTTAAATTTGTATATCCCTGATAGCTATAATTTGGGTCTATAGAAATTTCTT	4008
QY	2889	ACATGGGTTTGTGTTTACCCAGAAATGGAAGACCTGTACCCCTTTTGTGGTCTTTACTG	2948	QY	3969	TATACATTTCTAGATCAAGTCTTTGTCGGATATACGTTATGAGATATTTACACCTAGTCT	4028
DB	2939	ACATGGGTTTGTGTTTACCCAGAAATGGAAGACCTGTACCCCTTTTGTGGTCTTTACTG	2998	DB	4009	TATACATTTCTAGATCAAGTCTTTGTCGGATATATGTTATGAGATATTTACACCTAGTCT	4068
QY	2949	AGCTGGGTGGTCTGTTTGTAGCTTATTTAGAGTCCCTAGTTTCTCTTATTAAGAATA	3008	QY	4029	GTGGCTTCTGCTGTTTCTTATGCTTTTGTATGAATAGAGTTTAAATTTTGAAGGT	4088
DB	2999	AGCTGGGTGGTCTGTTTGTAGCTTATTTAGAGTCCCTAGTTTCTCTTATTAAGAATA	3058	DB	4069	GTGGCTTCTGCTGTTTCTTATGCTTTTGTATGAATAGAGTTTAAATTTTGAAGGT	4128
				QY	4089	CAAAATTTATTTTTTTTCTTTTGTGATATTTTTTCTCTCCAATTTTAAACCCCAAGATTTTCA	4148

[illegible]

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 168 Row: m Column: 13

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6996019. Location/Qualifiers

```

1. .4466
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="MGC:87362 IMAGE:30343504"
   /tissue_type="Placenta, normal"
   /clone_lib="NIH MGC_147"
   /lab_host="DH10B"
   /notes="Vector: pBluescript"
2. .4466
   /gene="NBS1"
   /note="synonyms: NIBRIN, AT-V1, AT-V2, ATV, NBS"
   /db_xref="LocusID:4683"
   /db_xref="MIM:602667"
3. .1516
   /gene="NBS1"
   /codon_start=1
   /product="NBS1 protein"
   /protein_id="AAH71590.1"
   /db_xref="GI:47938123"
   /db_xref="LocusID:4683"
   /db_xref="MIM:602667"
   /translation="NMKLPLPAAGPAGGPPYRLLTGTVYVGRKNCALITENDQSISRNS
   HAVLTAVRLEPLVACSQDIDTPEVLTKDMSKYGVFNVEEKMGQSFRTKLSGDGDTGQVGF
   GSKFRLEVPVLAQSDLVDSKTAALNOAILQLGFTVNNWTEECTHLVMVSVKVTIK
   TTICALICGRPIVVKPYFTFLKAVESKKQPPQIESFYPLDPEPSIGSKNVDLSGRQR
   KQIFGKPTFILNAQKHKLSSAVVFGGEARLITTEENBEHNFLAPQTCVVDGTGIT
   NTPGTLIDPCQKKIOSIMDLROQLRPIPEAIEGLAVIFMTTKNYCDQGHPSGTGLK
   TTTPTPSLSQGVSVDEKMLPSAPVNTTVADTESQADTDLSERPKEIKVSKMVEQK
   FRLMSQDAPTVKESCKTSSNNMSVNTLAKWRIPNYQLSPTKLPISINKSKDRASQQQ
   QTNSIRNYPQPSTKKKKGK"

```

ORIGIN

	Query Match	93.2%;	Score 4102;	DB 9;	Length 4466;
	Best Local Similarity	97.0%;	Pred. No. 0;		
	Matches 4279;	Conservative	7;	Mismatches 32;	Indels 92;
				Gaps	7;
QY	13	CGCGGTTGACGTCGCGCCCGCCAGCCCTGAGAGCCGACCGATGTGGAAACTGCTGCCCGC	72		
Db	67	CGCGGTTGACGTCGCGCCCGCCAGCCCTGAGAGCCGACCGATGTGGAAACTGCTGCCCGC	126		
QY	73	CGCGGCGCGGAGGAGGAGAACCATACAGACTTTTGTAGTGGCTTGTAGTACGTTGTTCG	132		
Db	127	CGCGGCGCGGAGGAGGAGAACCATACAGACTTTTGTAGTGGCTTGTAGTACGTTGTTCG	186		
QY	133	AAGGAAAAACTGTGCCATTCTTAATTGAAATGATCAGTCGATCAGCCGGAATCATGCTGT	192		
Db	187	AAGGAAAAACTGTGCCATTCTGATTGAAATGATCAGTCGATCAGCCGGAATCATGCTGT	246		
QY	193	GTTAACTGCTAACTTTTCTGTAAACCACTGTAGTCAACAGATGAAATCCCTGTATTGAC	252		
Db	247	GTTAACTGCTAACTTTTCTGTAAACCACTGTAGTCAACAGATGAAATCCCTGTATTGAC	306		
QY	253	ATTAAAGATAATCTTAAGTATGGTACCTTGTGTAATGAGGAAAAATGCAGAAATGGCTT	312		
Db	307	ATTAAAGATAATCTTAAGTATGGTACCTTGTGTAATGAGGAAAAATGCAGAAATGGCTT	366		
QY	313	TTCCCCGAACCTTTGAAGTCGGGGGATGGTATTACTTTTGGAGTGTGGAAGTAAATTCAG	372		
Db	367	TTCCCCGAACCTTTGAAGTCGGGGGATGGTATTACTTTTGGAGTGTGGAAGTAAATTCAG	426		
QY	373	AATAGAGTAGAGCCTTTGGTTGCATGCTCTTCTGTTTGTAGATGCTCTGGGAAACTGC	432		
Db	427	AATAGAGTAGAGCCTTTGGTTGCATGCTCTTCTGTTTGTAGATGCTCTGGGAAACTGC	486		

QY	433	TTTTAATCAAGCTATATTGCAACTTGGAGGATTTTACTGTAAACAAATTGGACAGAAATG	492	1512	GTCTCTTTTAGAACACAAACCTGCTACACCTCTATTGTGGAATAATAAGGACGAC	1571
Db	487	TTTTAAATCAAGCTATATTGCAACTTGGAGGATTTACTGTAAACAAATTGGACAGAAATG	546	1567	GTCTCTTTTAGAACACAAACCTGCTACACCTCTATTGTGGAATAATAAGGACGAC	1626
QY	493	CACTCACCTTGTTCATGGTATCAGTGAAGATTACCATTTAAACAAATATGCGACTCATTTTG	552	1572	ATCTATCTGAGAAATGAGCTGTGGACACAACTCAGACAATAACTTATTATTACAGATACAG	1631
Db	547	CACTCACCTTGTTCATGGTATCAGTGAAGATTACCATTTAAACAAATATGCGACTCATTTTG	606	1627	ATCTATCTGAGAAATGAGCTGTGGACACAACTCAGACAATAACTTATTATTACAGATACAG	1686
QY	553	TGACGCTCCAAATTGTAAAGCCAGAAATATTATTCTGAATTCCTGAAAGCGATTCAGTCCAA	612	1632	ATTTAAATCTATCTGAAATTTCTGCCAGTAAATCTCATGCTCGAGAAAGCTTAAGAT	1691
Db	607	TGACGCTCCAAATTGTAAAGCCAGAAATATTATTCTGAATTCCTGAAAGCGATTCAGTCCAA	666	1687	ATTTAAATCTATCTGAAATTTCTGCCAGTAAATCTCATGCTCGAGAAAGCTTAAGAT	1746
QY	613	GAAGCAGCTCCCAAAATTGAAAGTTTACCACCTCTTGATGAACCACTTATTTCGAAG	672	1692	CAAAATAAAAGGAGAAATGGATGATGTGGCCATAGAAGTGAAGTATTGGAAACAGTTAT	1751
Db	667	GAAGCAGCTCCCAAAATTGAAAGTTTACCACCTCTTGATGAACCACTTATTTCGAAG	726	1747	CAAAATAAAAGGAGAAATGGATGATGTGGCCATAGAAGTGAAGTATTGGAAACAGTTAT	1806
QY	673	TAAAAATGTTGATCTCTCAGGACGGCAGGAGAAAGAAACAAATCTTCAAAGGGGAAAAAT	732	1752	TCAAGGACACAAAAACAGAGTTAGAAATTTGATGTGAAATTTCAAAAAACAGGAGGAATG	1811
Db	727	TAAAAATGTTGATCTCTCAGGACGGCAGGAGAAAGAAACAAATCTTCAAAGGGGAAAAAT	786	1807	TCAAGGACACAAAAACAGAGTTAGAAATTTGATGTGAAATTTCAAAAAACAGGAGGAATG	1866
QY	733	TATATTTTGAATGCCAAACAGCATAAAGAAATTTGAGTTCGGCAGTTGTCTTTTGGAGGTGG	792	1812	TCAAATGTTAGAAAAAGGCCAAGATGGATATAGAAACAAATGACACTTTTCAGTGATGAAG	1871
Db	787	TATATTTTGAATGCCAAACAGCATAAAGAAATTTGAGTTCGGCAGTTGTCTTTTGGAGGTGG	846	1867	TCAAATGTTAGAAAAAGGCCAAGATGGATATAGAAACAAATGACACTTTTCAGTGATGAAG	1926
QY	793	GGAACTGAGTTGATACAGAGAGAAATGAAGAGAACATAATTTCTTTTGGCTCCGGG	852	1872	CAGTACCAAGAGTAGCAAAATATCTCAAGAAAAATGAAATTTGGGAAGAAACCTGGAATCA	1931
Db	847	GGAACTGAGTTGATACAGAGAGAAATGAAGAGAACATAATTTCTTTTGGCTCCGGG	906	1927	CAGTACCAAGAGTAGCAAAATATCTCAAGAAAAATGAAATTTGGGAAGAAACCTGGAATCA	1986
QY	853	AACGTGTGTTGTGATACAGGAAATAACAACTCACAGACCTTAAATTCCTGACTGTGAGAA	912	1932	AGGAAGACTACTATGGTCAGCTAAAGAAATATCTAAACAATGACAACTTCAGGATGATA	1991
Db	907	AACGTGTGTTGTGATACAGGAAATAACAACTCACAGACCTTAAATTCCTGACTGTGAGAA	966	1987	AGGAAGACTACTATGGTCAGCTAAAGAAATATCTAAACAATGACAACTTCAGGATGATA	2046
QY	913	GAAATGGAATCAGTCAATATGATATGCTCCAAAGGCAAGTCTTAGACCTATTCTCTGA	972	1992	GTGAGATGCTTCCAAAAAGCTTTATTGACCTGAATTTAGATCACTGGTGATTAAAAACT	2051
Db	967	GAAATGGAATCAGTCAATATGATATGCTCCAAAGGCAAGTCTTAGACCTATTCTCTGA	1026	2047	GTGAGATGCTTCCAAAAAGCTTTATTGACCTGAATTTAGATCACTGGTGATTAAAAACT	2106
QY	973	AGCAGAAATTTGATTCGGCGTGATTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGG	1032	2052	CTACTTCCAGAAATCCGCTCGCATAAATGATGATTTATGGTCAACTAAAAAATTTCAAGA	2111
Db	1027	AGCAGAAATTTGATTCGGCGTGATTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGG	1086	2107	CTACTTCCAGAAATCCATCTGCAATAATGATGATTTATGGTCAACTAAAAAATTTCAAGA	2166
QY	1033	CCATCCAGTACAGGATTTAAAGACAACTCCAGGACCAAGCTTTTCAAGAGCGGTGTC	1092	2112	AATTCAAAAAGGTCAATATCTCTGAGCAGGAAAACTTCCACACATCATTTGAGGATCAG	2171
Db	1087	CCATCCAGTACAGGATTTAAAGACAACTCCAGGACCAAGCTTTTCAAGAGCGGTGTC	1146	2167	AATTCAAAAAGGTCAATATCTCTGAGCAGGAAAACTTCCACACATCATTTGAGGATCAG	2226
QY	1093	AGTTGATGAAAACTTAATGCCAAAGCGCCAGTGGAACATAAACAATAGTACTGACAC	1152	2172	ATCTAATAGTCTCATCTCGAAGAAATACAGAACTAGAAGTGGCTTAAGCGAGGAAA	2231
Db	1147	AGTTGATGAAAACTTAATGCCAAAGCGCCAGTGGAACATAAACAATAGTACTGACAC	1206	2227	ATCTAATAGTCTCATCTCGAAGAAATACAGAACTAGAAGTGGCTTAAGCGAGGAAA	2286
QY	1153	AGAACTCAGACGACAGATACATGGGATTTGAGTGAAGGCCAAAGAAATCAAACTCTC	1212	2232	TGGAGGTACAAAATCAACATGCAAAAGAGAGTCTCTTCTGCTGATGATCTTTTTAGATACA	2291
Db	1207	AGAACTCAGACGACAGATACATGGGATTTGAGTGAAGGCCAAAGAAATCAAACTCTC	1266	2287	TGGAGGTACAAAATCAACATGCAAAAGAGAGTCTCTTCTGCTGATGATCTTTTTAGATACA	2346
QY	1213	CAAAATGGAACAAAAATTCAGAAATGCTTTTCAAGACGCAACCCACTGTAAAGAGTCCCTG	1272	2292	ATCTCTTATTTAAAAAGGAGAGATTAATCTGAGGATTTTAAAAAGAGCCATGGAATACTT	2351
Db	1267	CAAAATGGAACAAAAATTCAGAAATGCTTTTCAAGACGCAACCCACTGTAAAGAGTCCCTG	1326	2347	ATCTCTTATTTAAAAAGGAGAGATTAATCTGAGGATTTTAAAAAGAGCCATGGAATACTT	2406
QY	1273	CAAAAACAAGCTCTAATAATAATAGTATGGTATCAAAATATCTTTGGCTTAAGATGAATCCC	1332	2352	CCTAGTACGATCTACTTTCAGCCCAACAGGTTTATGAATATATAGTATAGTAGAGCGA	2411
Db	1327	CAAAAACAAGCTCTAATAATAATAGTATGGTATCAAAATATCTTTGGCTTAAGATGAATCCC	1386	2407	CCTAGTACGATCTACTTTCAGCCCAACAGGTTTATGAATATATAGTATAGTAGAGCGA	2466
QY	1333	AAACTATCAGCTTTTCAACCACTAAATTTGCCAGTATATAATAAGTAAAGATAGGCGCTTC	1392	2412	TTTAAAGTTACAAATGTTTATGGCTTAAATTTTAAATAAAATGCACAAAACTTTGATTC	2471
Db	1387	AAACTATCAGCTTTTCAACCACTAAATTTGCCAGTATATAATAAGTAAAGATAGGCGCTTC	1446	2467	TTTAAAGTTACAAATGTTTATGGCTTAAATTTTAAATAAAATGCACAAAACTTTGATTC	2526
QY	1393	TCAGCAGCAGCAGCAACCTCCATCAGAAACTACTTTTCAGCCGCTCTACC-AAAAAAGGG	1451	2472	TTTTGTATGTAACAAATGTTTGTCTGTTTTCAGGCTTTGTCAATTTGCAATCTTTTTTCAT	2531
Db	1447	TCAGCAGCAGCAGCAGCAACCTCCATCAGAACTACTTTTCAGCCGCTCTACC-AAAAAAGGG	1506	2527	TTTTGTATGTAACAAATGTTTGTCTGTTTTCAGGCTTTGTCAATTTGCAATCTTTTTTCAT	2586
QY	1452	AAAGGATGAGAAAAATCAAGAAATGTCTTATGCAAAATCAGCAAGAAATAGAAACGTCTTT	1511	2532	TTTTAAATGCTGTTTGTGTTTATTAATAGTTAATATAGTACACAGTTTCAAAATTTCTAAAT	2591
Db	1507	AAAGGATGAGAAAAATCAAGAAATGTCTTATGCAAAATCAGCAAGAAATAGAAACGTCTTT	1566	2587	TTTTAAATGCTGTTTGTGTTTATTAATAGTTAATATAGTACACAGTTTCAAAATTTCTAAAT	2646
				2592	ACGTAAGGTAAAGGACTAAAGTCAACCCTTCCACCATTGTCTTAGCTACT- - - - -	2640



3595 John Hopkins Court, San Diego, CA 92121, USA  
Tel: 18582029018  
Fax: 18582029020  
Email: abraun@sequenom.com  
Primer A: No primer sequence submitted  
Primer B: No primer sequence submitted  
STS size: 4423.

FEATURES  
source Location/Qualifiers  
1..4423  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="Human DNA (Sequenom)"  
<1..4423

Query Match	91.7%;	Score 4038.4;	DB 11;	Length 4423;
Best Local Similarity	96.6%;	Pred. No. 0;		
Matches 4270;	Conservative 8;	Mismatches 41;	Indels 103;	Gaps 11;
QY	12	GGCGGTTGCAGCTGGCCGCCAGCCCTGAGGAGCGGACCGGATGTGGAACTGCTGCCCG	71	
DB	4406	GGCGGTTGCAGCTGGCCGCC--AGCCTGAGGAGCGGACCGGATGTGGAACTGCTGCCCG	4349	
QY	72	CGCGGGCCCGGAGGAGGAAACCATACAGACTTTTGACTGGCGTTGAGTACGTTGTTG	131	
DB	4348	CGAGCG-CCGGCAGGAGGAGAACCATACAGACTTTTGACTGGCGTTGAGTACGTTGTTG	4290	
QY	132	GAGGAAAACTGTGCCATTCTAAATTGAAATGATCAGTCGATCAGCCGGAATCATGCTG	191	
DB	4289	GAGGAAAAAATGTGCCATTCTAAATTG-ACATGATCAGTCGATCAGCCGGAATCATGCTG	4231	
QY	192	TGTTAACTGCTAACTTTCTGTAAACCAACTGATCAACACAGATGAAATCCCTGTATTGA	251	
DB	4230	TGTTAACTGCTAACTTTCTGTAAACCAACTGATCAACACAGATGAAATCCCTGTATTGA	4171	
QY	252	CATTAAGATTAATCTTAAGTATGGTACCTTTGTTAATGAGGAAAAAATGCAGATGGCT	311	
DB	4170	CATTAAGATTAATCTTAAGTATGGTACCTTTGTTAATGAGGAAAAAATGCAGATGGCT	4111	
QY	312	TTTCCGAACTTTGAAATCGGGGATGGTATTACTTTTGGAGTGTGGAAGTAAATTC	371	
DB	4110	TTTCCGAACTTTGAAATCGGGGATGGTATTACTTTTGGAGTGTGGAAGTAAATTC	4051	
QY	372	GATAGATGATGAGCCTTGTGTGATGCTCTCTTGTGTTAGATGCTCTGGGAAACTG	431	
DB	4050	GATAGATGATGAGCCTTGTGTGATGCTCTCTTGTGTTAGATGCTCTGGGAAACTG	3991	
QY	432	CTTTAAATCAAGCTATATTGCAACTTGGAGGATTTACTGTAAACAAATTGGACAGAGAAT	491	
DB	3990	CTTTAAATCAAGCTATATTGCAACTTGGAGGATTTACTGTAAACAAATTGGACAGAGAAT	3931	
QY	492	GCACTCACCTTGTGATGATCAGTGAAGATTACCAATTAACAAATATGTGCACTCATTT	551	
DB	3930	GCACTCACCTTGTGATGATCAGTGAAGATTACCAATTAACAAATATGTGCACTCATTT	3871	
QY	552	GTGGAGCTCCAAATTGAAGCCAGATATTTTACTGAAATTCCTGGAAGCAGTTCACTCCA	611	
DB	3870	GTGGAGCTCCAAATTGAAGCCAGATATTTTACTGAAATTCCTGGAAGCAGTTCACTCCA	3811	
QY	612	AGAAGCAGCTCCACAAATTTGAAAGTTTATCCCACTCTTGATGAACCATCTATTGGAA	671	
DB	3810	AGAAGCAGCTCCACAAATTTGAAAGTTTATCCCACTCTTGATGAACCATCTATTGGAA	3751	
QY	672	GTAATAATGTTGATCTGTGAGGACGGCAGGAAAGAAAAAATAATCTTCAAGGGAAAAAT	731	
DB	3750	GTAATAATGTTGATCTGTGAGGACGGCAGGAAAGAAAAAATAATCTTCAAGGGAAAAAT	3691	
QY	732	TTATATTTTGAATGCCAAACAGCATAGAAATTTGAGTCCGAGTTGCTTTTGGAGGTG	791	
DB	3690	TTATATTTTGAATGCCAAACAGCATAGAAATTTGAGTCCGAGTTGCTTTTGGAGGTG	3631	
QY	792	GGGAAGCTAGTTGATAACAGAGAGATGAAGAGAACATAATTTCTTTTGGCTCCCG	851	

DB	3630	GGGAAGCTAGTTGATAACAGAGAGAATGAAGAAGACATAATTTCTTTTGGCTCCGG	3571	
QY	852	GAACTGTGTTGATACAGAGATAACAAACTACAGACCTTAATTTCTGACTCTCAGA	911	
DB	3570	GAACTGTGTTGATACAGAGATAACAAACTACAGACCTTAATTTCTGACTCTCAGA	3511	
QY	912	AGAAATGGAATTCAGTCAATTAATGATATGCTCCAAAGGCAGAGTCTTAGACCTATTCTCTG	971	
DB	3510	AGAAATGGAATTCAGTCAATTAATGATATGCTCCAAAGGCAGAGTCTTAGACCTATTCTCTG	3451	
QY	972	AAGCAGAAATTTGATTTGGCGGTGATTTTCATGACTACAAAGAAATTTCTGTGATCTCAGG	1031	
DB	3450	AAGCAGAAATTTGATTTGGCGGTGATTTTCATGACTACAAAGAAATTTCTGTGATCTCAGG	3391	
QY	1032	GCCATCCAGTACAGAGTTAAAGACAACACTCCAGGACCAAGCTTTTACAGAGCGGTG	1091	
DB	3390	GCCATCCAGTACAGAGTTAAAGACAACACTCCAGGACCAAGCTTTTACAGAGCGGTG	3331	
QY	1092	CAGTTGATGAAAAAATAATGCGCGCCCAAGTGAAACACTACAACTACAGTGTGCTGACA	1151	
DB	3330	CAGTTGATGAAAAAATAATGCGCGCCCAAGTGAAACACTACAACTACAGTGTGCTGACA	3271	
QY	1152	CAGAATCAGAGCAAGCAGATACATGGG-----ATTTGAGTGAAGGCGCAAAAGAAATCA	1205	
DB	3270	CAGAATCAGAGCAAGCAGATACATGGGATTTGATTTGAGTGAAGGCGCAAAAGAAATCA	3211	
QY	1206	AGTCTCCAAATGGAACAAATAATTCAGATGCTTTTCAAGACGCCACCTGTAAAGG	1265	
DB	3210	AGTCTCCAAATGGAACAAATAATTCAGATGCTTTTCAAGACGCCACCTGTAAAGG	3151	
QY	1266	AGTCTCCAAACAAAGCTCTAATAATAATAGTATGATGATCAAAATCTTTGGCTAAAGTGA	1325	
DB	3150	AGTCTCCAAACAAAGCTCTAATAATAATAGTATGATGATCAAAATCTTTGGCTAAAGTGA	3091	
QY	1326	GAAATCCCAAACTATCAGCTTTTCCAACTAAATTCGCAAGTATAATAAAGTAAAGATA	1385	
DB	3090	GAAATCCCAAACTATCAGCTTTTCCAACTAAATTCGCAAGTATAATAAAGTAAAGATA	3031	
QY	1386	GGGCTTCTCAGCAGCAGCAGCAACTCCATCAGAAATCTCTTTTCAGCGCTCTACCAAA	1445	
DB	3030	GGGCTTCTCAGCAGCAGCAGCAACTCCATCAGAAATCTCTTTTCAGCGCTCTACCAAA	2971	
QY	1446	AAAGGGAAGGGATCAAGAAATCAAGAAATGCTTTTCATGCAAAATCAGCAAGATAGAAA	1505	
DB	2970	AAAGGGAAGGGATCAAGAAATCAAGAAATGCTTTTCATGCAAAATCAGCAAGATAGAAA	2911	
QY	1506	CGTCTTGTCTCTTTTAGAAACAAACAACTGCTTACACCTCATCTTGGAAAAATAAGG	1565	
DB	2910	CGTCTTGTCTCTTTTAGAAACAAACAACTGCTTACACCTCATCTTGGAAAAATAAGG	2851	
QY	1566	AGCAGCATCTATCTCAGATGAGCCTGTGGACACAAACTCAGCAAAATTAATTTTACAG	1625	
DB	2850	AGCAGCATCTATCTCAGATGAGCCTGTGGACACAAACTCAGCAAAATTAATTTTACAG	2791	
QY	1626	ATACAGATTTAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGTCGAGAAAAAGC	1685	
DB	2790	ATACAGATTTAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGTCGAGAAAAAGC	2731	
QY	1686	TAAGATCAAAATAAAAAAGGGGAAATGGATGATGTGCCCATAGAAATGAAGTATTGGAAAC	1745	
DB	2730	TAAGATCAAAATAAAAAAGGGGAAATGGATGATGTGCCCATAGAAATGAAGTATTGGAAAC	2671	
QY	1746	AGTTATTCAGGACACAAACCCAGAGTTAGAAATTTGATGTGAAAGTTTCAAAAACAGGAGG	1805	
DB	2670	AGTTATTCAGGACACAAACCCAGAGTTAGAAATTTGATGTGAAAGTTTCAAAAACAGGAGG	2611	
QY	1806	AAAGATGCTCAATGTTAGAAAAAGCCCAAGGATGGATATAGAAAACAAATGACATTTTCAGTG	1865	
DB	2610	AAAGATGCTCAATGTTAGAAAAAGCCCAAGGATGGATATAGAAAACAAATGACATTTTCAGTG	2551	
QY	1866	ATGAAGCAGTACACAGAAAGTAGCAAAATATCTCAAGAAATTAATTTGGAGAAACCTG	1925	





```
variation      /replace="t"
2577
/genes="NBS1"
/insertion="0.37"
/replace="c"
2739
/genes="NBS1"
/insertion="0.01"
/replace="c"
2864
/genes="NBS1"
/insertion="0.36"
/replace="g"
2935..3324
/rpt_family="L1"
/rpt_type=dispersed
3000
/genes="NBS1"
/insertion="0.01"
/replace="g"
3051
/genes="NBS1"
/insertion="0.01"
/replace="a"
3151..3302
/rpt_family="Alu"
/rpt_type=dispersed
3226
/genes="NBS1"
/insertion="0.02"
/replace="g"
3338..3513
/rpt_family="L1"
/rpt_type=dispersed
3444
/genes="NBS1"
/insertion="0.04"
/replace="g"
3469
/genes="NBS1"
/insertion="0.37"
/replace="a"
3643..3646
/genes="NBS1"
/insertion="0.01"
/replace="n"
3816
/genes="NBS1"
/insertion="0.36"
/replace="a"
4018
/genes="NBS1"
/insertion="0.01"
/replace="g"
4182..4228
/rpt_family="Mariner"
/rpt_type=dispersed
4214
/genes="NBS1"
/insertion="0.02"
/replace="t"
4235..4281
/rpt_family="L2"
/rpt_type=dispersed
4353
/genes="NBS1"
/insertion="0.02"
/replace="a"
4426..4427
/genes="NBS1"
/insertion="0.01"
/replace="n"
4781..4782
```

```
variation      /genes="NBS1"
/insertion="0.01"
/replace="n"
4952
/genes="NBS1"
/insertion="0.35"
/replace="a"
5018

Query Match      42.3%; Score 1862.8; DB 9; Length 55134;
Best Local Similarity 94.1%; Pred. No. 0;
Matches 2024; Conservative 9; Mismatches 26; Indels 91; Gaps 6;

Qy 2278 TCCTTTTAGATACAATCCCTTTATTTAAAGGAGAGATAACTGAGGATTTTAAAGAAAG 2337
Db TCTTTGCAGATACAATCCCTTTATTTAAAGGAGAGATAACTGAGGATTTTAAAGAAAG 51057

Qy 2338 CCATGGAAAACTTCCTAGTAAAGCATCTCTTCAGGCCCAACAGGTTATATGAATATATA 2397
Db CCATGGAAAACTTCCTAGTAAAGCATCTCTTCAGGCCCAACAGGTTATATGAATATATA 51117

Qy 2398 GTGTATAGAGCGATTTAAAGTTACAAATGTTTTATGGCCTAAATTTATTAATAAAATGCA 2457
Db GTGTATAGAGCGATTTAAAGTTACAAATGTTTTATGGCCTAAATTTATTAATAAAATGCA 51177

Qy 2458 CAAACTTTGATCTTTTGTATGTAAACAATGTTGTCGTTTTGAGGCTTTGTCATTG 2517
Db CAAACTTTGATCTTTTGTATGTAAACAATGTTGTCGTTTTGAGGCTTTGTCATTG 51237

Qy 2518 CATCTTTTTCATTTTAAATGCTGTTTTGTTTTATTAATAGTTATATAGTCACAGTTC 2577
Db CATCTTTTTCATTTTAAATGCTGTTTTGTTTTATTAATAGTTATATAGTCACAGTTC 51297

Qy 2578 AAAATTCCTAAATCTAGCTAAGGTAAAGGACTAAAGTCAACCTTCCACCACTTGTCTAGCT 2637
Db AAAATTCCTAAATCTAGCTAAGGTAAAGTCAACCTTCCACCACTTGTCTAGCT 51356

Qy 2638 ACT----- 2640
Db ACTTGGTCCCTCAGAAAAAATTCATGATACTACTATTTCTTTATGAATCTTTCCAGGGAT 51416

Qy 2641 -----TATTTTAAATTAATTTCTACACAAATGATAGCTAAC 2678
Db TTTTGTAGTCTCTATTCCTATTTTAAATTAATTTCTACACAAATGATAGCTAAC 51476

Qy 2679 ATATGCAGTGTCTACACCTTGCTTTTACTTAGTAAGATTAAAAATTTATAGAAATATC 2738
Db ATATGCAGTGTCTACACCTTGCTTTTACTTAGT-AGATTAAAAATTTATAGAAATATC 51535

Qy 2739 AATATAATGTTTTTAATATTTTTTCTTTTCCATTTATGCTGTAGTCTTAACTCTGG 2798
Db AATATAATGTTTTTAATATTTTTTCTTTTCCATTTATGCTGTAGTCTTAACTCTGG 51595

Qy 2799 TGAATCCAAACAAATGGCTTCAGTGGTGAGATGTCACCTACATGTTATTTCTAGTACTAG 2858
Db TGAATCCAAACAAATGGCTTCAGTGGTGAGATGTCACCTACATGTTATTTCTAGTACTAG 51655

Qy 2859 AAATGAAGACCATGTGGAGACTTCATCAAAATGGGTTTAGTTTTTCCACAGAAATGAAA 2918
Db AAATGAAGACCATGTGGAGACTTCATCAAAATGGGTTTAGTTTTTCCACAGAAATGAAA 51715

Qy 2919 GACCTGTACCCCTTTTGGTGGTCTTAACTGAGCTGGGTGGTGTCTGTTTTGAGCTTATT 2978
Db GACCTGTACCCCTTTTGGTGGTCTTAACTGAGCTGGGTGGTGTCTGTTTTGAGCTTATT 51775

Qy 2979 TAGAGTCTAGTCTTCTACTTATTAAGTAGAAAATGGTGAGATGTTCTTTTCTTACC 3038
Db TAGAGTCTAGTCTTCTACTTATTAAGTAGAAAATGGTGAGATGTTCTTTTCTTACC 51835

Qy 3039 KTAAGGGAGATGGTAAAGAAACAAATGAATGCTTTTTTCAAACTTTTATGACAAGTATT 3098
Db KTAAGGGAGATGGTAAAGAAACAAATGAATGCTTTTTTCAAACTTTTATGACAAGTATT 51895
```

Qy	3099	TTCAAGTCTGTGTTCAAAAATATATTTCAATGTAACCTGTGATCAGCAAGAGGAGTTCCA	3158
Db	51896	TTCAAGTCTGTGTTCAAAAATATATTTCAATGTAACCTGTGATCAGCAAGAGGAGTTCCA	51955
Qy	3159	GTCAAGAGTCACCTACAACCTGATTTAGTTGTTTAGAGAAATGAGAAATGGAACAGTGAAGAAAT	3218
Db	51956	GTCAAGAGTCACCTACAACCTGATTTAGTTGTTTAGAGAAATGAGAAATGGAACAGTGAAGAAAT	52015
Qy	3219	GGAGGCCATATTTCCCATGACTTCCCTTGTATAACAGAAAGCAAACAGAAAGGGAACAAGAGCGCTG	3278
Db	52016	GGAGGCCATATTTCCCATGACTTCCCTTGTATAACAGAAAGCAAACAGAAAGGGAACAAGAGCGCTG	52075
Qy	3279	GCCTCTACATCACTCTCACCTTCCAAATCTTTGGGAAGTGCATCTACTTTGCCAGAACCAA	3338
Db	52076	GCCTCTACATCACTCTCACCTTCCAAATCTTTGGGAAGTGCATCTACTTTGCCAGAACCAA	52135
Qy	3339	ATTTAACTTACTTCCAAGTCTTCGCTGCTTGCAGGTGGAACTCCAGCTGCAAGGAGATTAG	3398
Db	52136	ATTTAACTTACTTCCAAGTCTTCGCTGCTTGCAGGTGGAACTCCAGCTGCAAGGAGATTAG	52195
Qy	3399	GGAAATCAAGGCTCTTTTTTTTAAAAGCTTCTCAGCCCTTCTGAGGAACAGAAATTCGGTGA	3458
Db	52196	GGAAATCAAGGCTCTTTTTTTTAAAAGCTTCTCAGCCCTTCTGAGGAACAGAAATTCGGTGA	52255
Qy	3459	GCCAACTCTGCAAAATTTCTACTACAGGCATTTGAGACCAAGTTAGATTATTTGAAATATTTATAGA	3518
Db	52256	GCCAACTCTGCAAAATTTCTACTACAGGCATTTGAGACCAAGTTAGATTATTTGAAATATTTATAGA	52315
Qy	3519	GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAAACATGGGATACCTTT	3578
Db	52316	GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAAACATGGGATACCTTT	52375
Qy	3579	AGAAGTAGAAATTCGACAGGGCATATTAAGTTGATGAAATGGAGTCATTTGAGTCTCTTTAATA	3638
Db	52376	AGAAGTAGAAATTCGACAGGGCATATTAAGTTGATGAAATGGAGTCATTTGAGTCTCTTTAATA	52435
Qy	3639	GCCATGTATCATATAATTACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTA	3698
Db	52436	GCCATGTATCATATAATTACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTA	52495
Qy	3699	AGGAATATAATGGACAGATTAAATATTGTTTCTGTCATGCGCACAAATCCCTTTCTTAAGGA	3758
Db	52496	AGGAATATAATGGACAGATTAAATATTGTTTCTGTCATGCGCACAAATCCCTTTCTTAAGGA	52555
Qy	3759	AGACTGCCCTACTATACGAGTCTTTTATATTGTCAAATTTATGAATATATGAATCAGGAG	3818
Db	52556	AGACTGCCCTACTATACGAGTCTTTTATATTGTCAAATTTATGAATATATGAATCAGGAG	52614
Qy	3819	TTCTGGTACCTCTCTGTCTTTACAAATATTTGGGTGTTGTCCAGTATTTTTCCCTTTTTTAAC	3878
Db	52615	TTCTGGTACCTCTCTGTCTTTACAAATATTTGGGTGTTGTCCAGTATTTTTCCCTTTTTTAAC	52670
Qy	3879	QMTTCCAAATTCGGGTGTGATGGAGTGTTCATATTTGGGTTTAAATTTGTAATATCCCT	3938
Db	52671	CCAATCCAAATTCGGGTGTGATGGAGTGTTCATATTTGGGTTTAAATTTGTAATATCCCT	52725
Qy	3939	GATAGCTATAATTTGGGTGCATAGAAATTTCTTTATACATTTCTAGATGCAAGTCTCTGCGG	3998
Db	52726	GATAGCTATAATTTGGGTGCATAGAAATTTCTTTATACATTTCTAGATGCAAGTCTCTGCGG	52785
Qy	3999	ATATACGTTATTTAGATATTACACTAGTCTGTGGCTTGACTGTGTTTTCTTTATGCTCTTTTG	4058
Db	52786	ATATATGTTATTTAGATATTACACTAGTCTGTGGCTTGACTGTGTTTTCTTTATGCTCTTTTG	52845
Qy	4059	ATGAATAGAAAGTTTTAAATTTTGAACAGGTCAAATTTATTTTTTCTTTTTGTTTGTATATT	4118
Db	52846	ATGAATAGAAAGTTTTAAATTTTGAACAGGTCAAATTTATTTTTTCTTTTTGTTTGTATATT	52905
Qy	4119	TTTTTCTCTCCAAATTTTAAACCCAGATTTTCCAGATATTTCTGCTCTATTATATAAATTTATA	4178
Db	52906	TTTTTCTCTCCAAATTTTAAACCCAGATTTTCCAGATATTTCTGCTCTATTATATAAATTTATA	52965
Qy	4179	TTTTTTATATTTGTGATCTACCTTTGAAATTTGATATGTATGTTGTGAATTTATGGATCAGGGTT	4238

Db 52966 TTTTATATTGTCATACCTTGAATTCATATGATGTTGTGAATTCAGGTT 53025  
 QY 4239 CTTTTCCTCCCATCAAGTATCCAGTCATTTGTAACACTGTTTATCGAAGAATTATCC 4298  
 Db 53026 CTTTTCCTCCCATCAAGTATCCAGTCATTTGTAACACTGTTTATCGAAGAATTATCC 53085  
 QY 4299 TTTTCCTCATTAATACCTTGCCAAATAGTAAATAATCAATTAACCATRM 4348  
 Db 53086 TTTTCCTCATTAATACCTTGCCAAATAGTAAATAATCAATTAACCATAA 53135  
 RESULT 8  
 LOCUS AB013139 56500 bp DNA linear PRI 25-DEC-2002  
 DEFINITION Homo sapiens gene for NBS1, complete cds.  
 ACCESSION AB013139  
 VERSION AB013139.1 GI:3169124  
 KEYWORDS NBS1; Nijmegen breakage syndrome.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Matsura, S., Tauchi, H., Nakamura, A., Kondo, N., Sakamoto, S.,  
 Endo, S., Smeets, D., Solder, B., Belohradsky, B. H., Kaloustian, V. M.,  
 Oshimura, M., Iomura, M., Nakamura, Y. and Komatsu, K.  
 TITLE Positional cloning of the gene for Nijmegen breakage syndrome  
 JOURNAL Nat. Genet. 19 (2), 179-181 (1998)  
 MEDLINE 98282099  
 PUBMED 9620777  
 REFERENCE 2 (bases 1 to 56500)  
 AUTHORS Matsura, S., Tauchi, H. and Komatsu, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-APR-1998) Shinya Matsura, Hiroshima University,  
 Research Institute for Radiation Biology and Medicine, Department  
 of Radiation Biology; Kasumi 1-2-3, Minami-ku, Hiroshima, Hiroshima  
 734-8553, Japan (E-mail: shinya@ipc.hiroshima-u.ac.jp,  
 Tel: 81-82-257-5811, Fax: 81-82-256-7101)  
 COMMENT Sequence updated (26-May-1998).  
 FEATURES  
 source  
 /organism="Homo sapiens"  
 /mol\_type="Genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /map="8q21"  
 /sub\_clone="RG255A7"  
 4790..4848  
 /number=1  
 4790..4811  
 join(4812..4848, 6518..6651, 7850..7998, 8480..8639,  
 11050..11153, 18083..18200, 18816..19009, 24866..24963,  
 30519..30648, 33818..34090, 35682..36129, 41481..41549,  
 43078..43233, 46007..46120, 52298..52347, 53761..53791)  
 /codon\_start=1  
 /product="NBS1"  
 /protein\_id="BAA28616.1"  
 /db\_xref="GI:3169125"  
 /translation="MWKLDPAGAGPAGGEPYRLLTGVEVVGKNCACILINDQSISRN  
 HAVLTANFVTLSDTDELPVLTLDKNSKYGFVNEEKQNGFSRLTKSGDITPGV  
 GSKFTEYPLVACSLDVGKALNQALQLGFTVTNNWTECTHLVMVSVKVIK  
 TICALLIGPTIVKPEYTFPLKAVESKQPOIESFYPLDPEPSIGSKNVLSGRQR  
 KOIPLKGTITFLNAQHKLLSAVVGFGGEARLITEENEEHNFPLAPTGCVDTGIT  
 NSQTLLPCQCKWIGSIMDMQROGRPIPEAEISGLAVIPMTIKNYCDPQGHPTGLK  
 TTPTPGLSCQGVSDKMLSPAVNTTYVADTESEQADTVADTSERKPEIKVSRMEQ  
 FRMLSDQAPTVKESCTSSNNKSMVNTLAKRIIPNYQLSPTKGPSINKSKDRASQQ  
 QTNISIRNYFPQPTKKREEDNEQMSCKSAIETSCSLLEQTQPATPSLWKNLEQL  
 SENEPYDSDNNLFTDYLKIVKNSAKSHAAKLRNKRMDVDAIEDVLEQL  
 FKDTPELIDIVKQKQEDVNRKPRMDIETNDTFSDAEVPSSKISQNEIGKKR  
 ELKEDSLMSAKEISNNKLDQDSEMLPKLLLTFRSLVIKNTSRNPSGINDDYGOL  
 KNPKFKPKYVTFGAKGLPHLIGSDLIHAHHRNTELEBWLKQMEVQNQHAKESRLA  
 DDLFRFKNYVTKRRR"

intron	4849..6517	Db		53812	CCATGGAAAACCTTCTAGTAAGCATCTACTTCAGGCCAACAAAGGTTATATGTAATATATA	53871
exon	/number=1	Qy		2398	GTGTATAGAAAGCGATTAAAGTTTACAAATGTTTATGCGCTAAATTTATTAATAAATGCA	2457
intron	/number=2	Db		53872	GTGTATAGAAAGCGATTAAAGTTTACAAATGTTTATGCGCTAAATTTATTAATAAATGCA	53931
exon	/number=2	Qy		2458	CAAAACCTTTGATCTTTTGTATGTAACAATTTGTTTCTGTTTTCAGGCTTTGTCAATG	2517
intron	/number=3	Db		53932	CAAAACCTTTGATCTTTTGTATGTAACAATTTGTTTCTGTTTTCAGGCTTTGTCAATG	53991
exon	/number=3	Qy		2518	CATCTTTTTCATTTTAAATGTTTGTGTTTATTAATAGTAAATATAGTACAGTTC	2577
intron	/number=4	Db		53992	CATCTTTTTCATTTTAAATGTTTGTGTTTATTAATAGTAAATATAGTACAGTTC	54051
exon	/number=5	Qy		2578	AAATTTCTAAATRTACGTAAAGGTAAAGGACTAAAGTCAACCTTCCACCACTTGTCTAGCT	2637
intron	/number=5	Db		54052	AAATTTCTAAATRTACGTAAAGGTAAAGTCAACCTTCCACCACTTGTCTAGCT	54110
exon	/number=6	Qy		2638	ACT-----	2640
intron	/number=6	Db		54111	ACTTGGTTCCTCCCTCAGAAAAAATTCATGATCTCATTTCTTATGAATCTTTCACGGGAT	54170
exon	/number=7	Qy		2641	-----TATTTTAAATAATTTCTTACACAAATGATAGCATAAC	2678
intron	/number=7	Db		54171	TTTGTAGTCTTATTCAAATTCCTATTTTAAATAATTTCTTACACAAATGATAGCATAAC	54230
exon	/number=8	Qy		2679	ATATGCAGTGTCTACACCTTGCTTTTACCTTTAGTAAGATTAAAAATTATAGGAATATC	2738
intron	/number=8	Db		54231	ATATGCAGTGTCTACACCTTGCTTTTACCTTAGT-AGATTAANAATTTATAGGAATATC	54289
exon	/number=9	Qy		2739	AATATAATGTTTAAATATTTTCTTTTCCATTTATGCTGTAGTCTTACCTAAACTCTGG	2798
intron	/number=9	Db		54290	AATATAATGTTTAAATATTTTCTTTTCCATTTATGCTGTAGTCTTACCTAAACTCTGG	54349
exon	/number=10	Qy		2799	TGATCCAAACAAATGGCTTCAGTGTGCAGATGTCCACTACATGTTATCTAGTACTAG	2858
intron	/number=10	Db		54350	TGATCCAAACAAATGGCTTCAGTGTGCAGATGTCCACTACATGTTATCTTGTAGTACTAG	54409
exon	/number=11	Qy		2859	AAACTGAAGACCATGTGGAGACTTCATCAAAACATGGGTTTAGTTTTTACCAGAAATGAAA	2918
intron	/number=11	Db		54410	AAACTGAAGACCATGTGGAGACTTCATCAAAACATGGGTTTAGTTTTTACCAGAAATGAAA	54469
exon	/number=12	Qy		2919	GACCTGTACCCCTTTTGGTGGTCTTAACTGAGCTGGGTGGGTGTCTGTTTGGAGCTTAT	2978
intron	/number=12	Db		54470	GACCTGTACCCCTTTTGGTGGTCTTAACTGAGCTGGGTGGGTGTCTGTTTGGAGCTTAT	54529
exon	/number=13	Qy		2979	TAGAGTCTAGTTTCTTACTTATAAAGTAGAAATGGTGCAGATGTTTCTTTTCTTACC	3038
intron	/number=13	Db		54530	TAGAGTCTAGTTTCTTACTTATAAAGTAGAAATGGTGCAGATGTTTCTTTTCTTACC	54589
exon	/number=14	Qy		3039	KTAAAGGGAGATGGTAAAGAAACAAATGAATGTCTTTTTCAAACTTTTATTCACAAGTGATT	3098
intron	/number=14	Db		54590	TTAAAGGGAGATGGTAAAGAAACAAATGAATGTCTTTTTCAAACTTTTATTCACAAGTGATT	54649
exon	/number=15	Qy		3099	TTCAAGTCTGTGTTCAAAAATATATTCATGTACTGTGATCCAGCAAGAGGAGTTCCA	3158
intron	/number=15	Db		54650	TTCAAGTCTGTGTTCAAAAATATATTCATGTACTGTGATCCAGCAAGAGGAGTTCCA	54709
exon	/number=16	Qy		3159	GTCAAGAGTCACCTACAACTGATTTGTTTGTAGAGTAAGAAATGGAACAGTCAGGAAT	3218
intron	/number=16	Db		54710	GTCAAGAGTCACCTACAACTGATTTGTTTGTAGAGTAAGAAATGGAACAGTCAGGAAT	54769
exon	/number=17	Qy		3219	GGAGGCCATATTTCCATGACTTCCCTTGTAAACAGAAAGCAACAGAGGGACAGAGGCTG	3278
intron	/number=17	Db		54770	GGAGGCCATATTTCCATGACTTCCCTTGTAAACAGAAAGCAACAGAGGGACAGAGGCTG	54829
exon	/number=18	Qy		3279	GCCTCTACATCACTCTCACCTTCCAAATCTTGTGGAAGTGCATCTACTTTGCCAGAACCAA	3338
intron	/number=18	Db		54830	GCCTCTACATCACTCTCACCTTCCAAATCTTGTGGAAGTGCATCTACTTTGCCAGAACCAA	54889
exon	/number=19	Qy		3339	ATTAACTTACTTCCAGTTTCTGGCTGCTTGGAGGTGGAACTCCAGCTGCAAGGGAGTTAG	3398

Query Match 42.3%; Score 1862.8; DB 9; Length 56500;  
Best Local Similarity 94.1%; Pred. No. 0;  
Matches 2024; Conservative 9; Mismatches 26; Indels 91; Gaps 6;  
Qy 2278 TCTTTTGTAGTACAACTCTTATTTTAAAGGAGAGTAACTGAGGATTTTAAAGAGAG 2337  
Db 53752 TCTTTCAGTACAACTCTTATTTTAAAGGAGAGTAACTGAGGATTTTAAAGAGAG 53811  
Qy 2338 CCATGGAAAACCTTCTAGTAAGCATCTACTTCAGGCCCAACAAAGGTTATATGAATATATA 2397

Db	54890	ATTAACCTTACTTCCAAAGTCTTCGCTGCTGCGAGGTGGAACTCCAGCTGCAAGGGAGTTAG	54949
Qy	3399	GGAAATGAAGGTCTTTTAAAAAGCTTCTCAGCCCTTCCTAGGGAACAGAAATGGGTGA	3458
Db	54950	GGAAATGAAGGTCTTTTAAAAAGCTTCTCAGCCCTTCCTAGGGAACAGAAATGGGTGA	55009
Qy	3459	GCCAAATCTGCAATTTCTACTACAGGCATTTAGATTAATTTGAAATATTATATAGA	3518
Db	55010	GCCAAATCTGCAATTTCTACTACAGGCATTTAGATTAATTTGAAATATTATATAGA	55069
-Qy	3519	GAGTTATGAACACTTAAATTTATCATAGTGGTATGACATTTGGATAGAACATGGGATACTTT	3578
Db	55070	GAGTTATGAACACTTAAATTTATCATAGTGGTATGACATTTGGATAGAACATGGGATACTTT	55129
Qy	3579	AGAAATGAATTTGACAGGGCATATAGTTAGTGAATGGAGTCAATTTGAGTCTTAAATA	3638
Db	55130	AGAAATGAATTTGACAGGGCATATAGTTAGTGAATGGAGTCAATTTGAGTCTTAAATA	55189
Qy	3639	GCCATGTATCATAAATTACCAAGTGAAGCTGGTGGAAATATGATGATCTCCATTTTACAGTTA	3698
Db	55190	GCCATGTATCATAAATTACCAAGTGAAGCTGGTGGAAATATGATGATCTCCATTTTACAGTTA	55249
Qy	3699	AGAAATGAATTTGACAGGGCATATAGTTAGTGAATGGAGTCAATTTGAGTCTTAAATA	3758
Db	55250	AGAAATGAATTTGACAGGGCATATAGTTAGTGAATGGAGTCAATTTGAGTCTTAAATA	55309
Qy	3759	AGCTGCGCTTACTATAGCAGTCTTTTATATTTGTCATTTTGAATTAATGAATGAATGAGAG	3818
Db	55310	AGCTGCGCTTACTATAGCAGTCTTTTATATTTGTCATTTTGAATTAATGAATGAATGAGAG	55368
Qy	3819	TTCTGTTACCTCTCTCTTTTACAAATTAATTTGGTGGTGGTCCAGTATTTTCCCTTTTAAAC	3878
Db	55369	TTCTGTTACCTCTCTCTTTTACAAATTAATTTGGTGGTGGTCCAGTATTTTCCCTTTTAAAC	55424
Qy	3879	CMTTCCCAATTCGGGTGTAGTGGAGTTTCCATTTGGGTTTAAATTTGATATATCCCT	3938
Db	55425	CCATTTCCAAATTCGGGTGTAGTGGAGTTTCCATTTGGGTTTAAATTTGATATATCCCT	55479
Qy	3939	GATAGCTATTAATTTGGGTGATAGAAATTTCTTTATACATTTCTAGATGCAAGTCTCTGCGG	3998
Db	55480	GATAGCTATTAATTTGGGTGATAGAAATTTCTTTATACATTTCTAGATGCAAGTCTCTGCGG	55539
Qy	3999	ATATACGATTTAGATATATTACACCTAGTCTGTGGCTTGACTGTTTCTTTATGCTCTTTG	4058
Db	55540	ATATACGATTTAGATATATTACACCTAGTCTGTGGCTTGACTGTTTCTTTATGCTCTTTG	55599
Qy	4059	ATGAATGAGAGTTTAAATTTTACAGAGTCAAAATTTATTTTCTTTTGTGTTGATATTT	4118
Db	55600	ATGAATGAGAGTTTAAATTTTACAGAGTCAAAATTTATTTTCTTTTGTGTTGATATTT	55659
Qy	4119	TTTTTCTCTCAATTTTAAATTTTACAGAGTCAAAATTTATTTTCTTTTGTGTTGATATTT	4178
Db	55660	TTTTTCTCTCAATTTTAAATTTTACAGAGTCAAAATTTATTTTCTTTTGTGTTGATATTT	55719
Qy	4179	TTTTTATATTGTTGATCTACCTTTCAATTTGATATGATGTTGTTGAAATTTGATGAGGTTT	4238
Db	55720	TTTTTATATTGTTGATCTACCTTTCAATTTGATATGATGTTGTTGAAATTTGATGAGGTTT	55779
Qy	4239	CTTTTTTCCCCCATACAGATTTCCAGTCAATTTGATGATGATGATGATGATGATGATGATGAT	4298
Db	55780	CTTTTTTCCCCCATACAGATTTCCAGTCAATTTGATGATGATGATGATGATGATGATGATGAT	55839
Qy	4299	TTTTTCTCATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT	4348
Db	55840	TTTTTCTCATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT	55899
RESULT 9			
LOCUS	AF069291/c	184919 bp	DNA linear PRI 02-OCT-1998
DEFINITION	Homo sapiens BAC clone 255A7 from 8q21 containing NBS1 gene, complete sequence.		
ACCESSION	AF069291		

```
exon      /number=3
73635..73826
/genes="hr41"
/number=4
80160..80251
/genes="hr41"
/number=5
83583..86947
/genes="hr41"
/number=6
complement (94512..143663)
/genes="NBS1"
/notes="Gene for NiJmegen breakage syndrome"
complement (join(94512..94692,96106..96155,102333..102446,
105220..105375,106904..106972,112324..112771,
114363..114635,117805..117934,123490..123587,
129444..129637,130253..130370,137300..137403,
139814..139973,140455..140603,141802..141935,
143605..143663))
/genes="NBS1"
/product="nibrin"
complement (94512..94692)
/number=16
complement (join(94662..94692,96106..96155,102333..102446,
105220..105375,106904..106972,112324..112771,
114363..114635,117805..117934,123490..123587,
129444..129637,130253..130370,137300..137403,
139814..139973,140455..140603,141802..141935,
143605..143641))
/genes="NBS1"
/notes="similar to yeast Xrs2"
/codon_start=1
/product="nibrin"
/db_xref="GI:3687830"
/translation="MKKLLPAAAGPAGGEPYRLLTGVYEVVGRKNCAILLNDQSIERN
HAVLTANFVNLASQDEIPVTLKDNSKYGTGVNBERQNGFSRTLKSGDGIITGFGV
GSKFRLEYPLVNLSCSDVSGTALNQAILQGGFTVNNWTEETHLVMVSVKVTIK
TICALICGRPIVKNPEYFTFLKAVESKKQPPQIESYPPPLDEPISGKNVDLSGRGR
KQIFKGTTFILNAKOHKLLSAVFGGGEARLITEENEEHNFLIACPTCVDTGIT
NSQTLIPDCQKWIQSIMDMLQRLPIPEAIEGLAVIPMTKNYCDQHPHSGLK
TTTPGSLQGVSVDEKLPMSAVNTTIVADTESQADTWLSEPKRPIKVSOMEQK
FRLMSQADPTVESCTSSNNNTTTLAKRIPNYQLSPKLPSINKSKORASQQQ
QNSIRNYQPSTKREDEENQMSCKSARIETSCSLEQTQPATPISLKNKEQHL
SENEPVDVTDNNLFTDOLKSIKNASAKSHAAEKLRSNKKREMDVAIIDEVLEQL
FKDTKPELIDIVKQKQEDVNVKPRMDIETNDTFSDEAVPESSKIQENIEGKR
ELKEDLSLAKESINNDLQDSEMLPKLLLTETFSRLVINKSTSRNPSGINDDYGQL
KNPKFKKVTYPGAGKLPHIIGGSDLIHAHARKNTELEWLKQEMEVQNQHAKESLSA
DDLFRNYPYLKRRR"
complement (96106..96155)
/genes="NBS1"
/number=15
complement (102333..102446)
/genes="NBS1"
/number=14
complement (105220..105375)
/genes="NBS1"
/number=13
complement (106904..106972)
/genes="NBS1"
/number=12
complement (112324..112771)
/genes="NBS1"
/number=11
complement (114363..114635)
/genes="NBS1"
/number=10
complement (117805..117934)
/genes="NBS1"
/number=9
complement (123490..123587)
/genes="NBS1"
```

```
exon      /number=8
complement (129444..129637)
/genes="NBS1"
/number=7
complement (130253..130370)
/genes="NBS1"
/number=6
complement (137300..137403)
/genes="NBS1"
/number=5
complement (139814..139973)
/genes="NBS1"
/number=4
complement (140455..140603)
/genes="NBS1"
/number=3
complement (141802..141935)
/genes="NBS1"
/number=2
complement (143605..143663)
/genes="NBS1"
/number=1
160500..>180136
/genes="DECR"
join(160500..160641,176204..176406,177989..178046,
178166..178252,179989..>180136)
/genes="DECR"
/product="2,4-dienoyl-CoA reductase"
160500..160641
/genes="DECR"
/number=1
join(160573..160641,176204..176406,177989..178046,
178166..178252,179989..>180136)
/genes="DECR"
/codon_start=1
/product="2,4-dienoyl-CoA reductase"
/protein_id="AAC62232.1"
/db_xref="GI:3687831"
/translation="MKLIPARVFTLGSRLPCGLAPRRFSGYGTIKLYQNTALQSKFF
SPLOKAMPPNSFOCKVAFITGGGTGLGKMTLLSSLGAQCQVIASRKMVDLKATAEQ
ISSQNGKLVHAIQCVRDPDMVQNTVSELIKVAGHPNIVINNAAGNFISPTERLSPNA
WKTIIDVLNGTAFVTLRIGKQLIKAQK"
176204..176406
/genes="DECR"
/number=2
177989..178046
/genes="DECR"
/number=3
178166..178252
/genes="DECR"

Query Match      42.3%; Score 1862.8; DB 9; Length 184919;
Best Local Similarity 94.1%; Pred. No. 0;
Matches 2024; Conservative 9; Mismatches 26; Indels 91; Gaps 6;

QY  2278  TCTTTTGTAGATACAACTCCTTATTATTAAGAGAGAGATAAATCTCAGGATTTTAAAGAAAG 2337
      |||||
DB  94701  TCTTTCAGATACAACTCCTTATTATTAAGAGAGAGATAAATCTCAGGATTTTAAAGAAAG 94642

QY  2338  CCATGGAAAAACTTCTTAGTAGCATCTTCTCAGGCCCAACAGGTATATATGAAATATA 2397
      |||||
DB  94641  CCATGGAAAAACTTCTTAGTAGCATCTTCTCAGGCCCAACAGGTATATATGAAATATA 94582

QY  2398  GTGTATAGAGCGATTAAAGTTTACAAATGTTTTATGCCTTAAATTTATTAAATAAATGCA 2457
      |||||
DB  94581  GTGTATAGAGCGATTAAAGTTTACAAATGTTTTATGCCTTAAATTTATTAAATAAATGCA 94522

QY  2458  CAAAACCTTTGATCTTTTGTATGTAACAATTTGTTGTCYCTGTTTTCAGGCTTTGTCAATG 2517
      |||||
DB  94521  CAAAACCTTTGATCTTTTGTATGTAACAATTTGTTGTCYCTGTTTTCAGGCTTTGTCAATG 94462

QY  2518  CATCTTTTTTTCATTTTAAATGCTTTGTTTATTAAATAGTTTAAATATAGTCACAGTTC 2577
      |||||
```

Db	94461	CATCTTTTTCATTTTTTAAATGTGTTTTGTTTTTAAATAGTTAATATAGTCACAGTTC	94402
Qy	2578	AAATTTCTAAATRTAGCTAAGGTAAAGGACTAAAGTCAACCTTCACCAATTTGCTCTAGCT	2637
Db	94401	AAATTTCTAAATGTACGTAAAGGTAAA - GACTAAGTCAACCTTCACCAATTTGCTCTAGCT	94343
Qy	2638	ACT-----	2640
Db	94342	ACTTGGTCCCTCAGAAAAAAATTCATGATCTCAATCTCTTATGAAATCTTTCCAGGGAT	94283
Qy	2641	-----TATTTTAAATAATTTCTCTACACAAATGATAGCATAC	2678
Db	94282	TTTTGAGTCCTATTCAAATTCCTATTTTAAATAATTTCTACACAAATGATAGCATAC	94223
Qy	2679	ATATGACAGTGTCTACACCTTGCTTTTATCTTAGTAAAGTTAAATAATATAGGAATATC	2738
Db	94222	ATATGACAGTGTCTACACCTTGCTTTTATCTTAGT - AGAATAAATAATATAGGAATATC	94164
Qy	2739	AAATATAATGTTTTTAATATTTTTTTTCCATTATGCTGTAGTCTTACCTAAACTCTCG	2798
Db	94163	AAATATAATGTTTTTAATATTTTTTTTCTTTCCATTATGCTGTAGTCTTACCTAAACTCTCG	94104
Qy	2799	TGATCCAAACAAATGGCTTCAGTGTGCAGATGTCACCTACATGTTATTTCTAGTACTAG	2858
Db	94103	TGATCCAAACAAATGGCTTCAGTGTGCAGATGTCACCTACATGTTATTTCTAGTACTAG	94044
Qy	2859	AACTGAAGACCATGTGGAGACTTCAATCAACATGGGTTAGTTTTTACCAGAAATGGAAA	2918
Db	94043	AAACTGAAGACCATGTGGAGACTTCAATCAACATGGGTTAGTTTTTACCAGAAATGGAAA	93984
Qy	2919	GACCTGTACCCCTTTTGGTGGTCTTACGTAGCTGGTGGTCTGTGTTTTTGGAGCTTAT	2978
Db	93983	GACCTGTACCCCTTTTGGTGGTCTTACGTAGCTGGTGGTCTGTGTTTTTGGAGCTTAT	93924
Qy	2979	TAGAGTCTAGTTTTCTCTACTATATAAGTAGAAATGGTGAGATTTGTTTTCTTTCTACC	3038
Db	93923	TAGAGTCTAGTTTTCTCTACTATATAAGTAGAAATGGTGAGATTTGTTTTCTTTCTACC	93864
Qy	3039	KTAAAGGGAGATGGTAAAGAAACATGAATGTCTTTTTTCAACTTTATTTGACAGTGAT	3098
Db	93863	TTAAAGGGAGATGGTAAAGAAACATGAATGTCTTTTTTCAAACTTTATTTGACAAGTGAT	93804
Qy	3099	TTCAAGTCTGTGTTCAAAATAATATTATCATGTCTGTGATCCAGCAAGAGGAGTTCCA	3158
Db	93803	TTCAAGTCTGTGTTCAAAATAATATTATCATGTCTGTGATCCAGCAAGAGGAGTTCCA	93744
Qy	3159	GTCAAGAGTCACTACAACCTGATTAGTTGTTAGAGAAATGAGAAATGGAAACAGTGAGGAAT	3218
Db	93743	GTCAAGAGTCACTACAACCTGATTAGTTGTTAGAGAAATGAGAAATGGAAACAGTGAGGAAT	93684
Qy	3219	GGAGGCCATATTTCCATGACTTCCCTTTGTAACAGAAAGCAACAGAAAGGACAGAGGCTG	3278
Db	93683	GGAGGCCATATTTCCATGACTTCCCTTTGTAACAGAAAGCAACAGAAAGGACAGAGGCTG	93624
Qy	3279	GCCTCTACATCACTCTCACTTCCAAATCTTGTGGAAGTGCATCTACTTGCAGAACCAA	3338
Db	93623	GCCTCTACATCACTCTCACTTCCAAATCTTGTGGAAGTGCATCTACTTGCAGAACCAA	93564
Qy	3339	ATTAACTTACTTCCAAAGTCTTGCTGCTTGCAGGTGGAACTCCAGCTGCAAGGGAGTTAG	3398
Db	93563	ATTAACTTACTTCCAAAGTCTTGCTGCTTGCAGGTGGAACTCCAGCTGCAAGGGAGTTAG	93504
Qy	3399	GGAAATGAAGGCTTTTTTAAAGCTTCTCAGCCTTCTAGGGAAACAGAAATTTGGGTGA	3458
Db	93503	GGAAATGAAGGCTTTTTTAAAGCTTCTCAGCCTTCTAGGGAAACAGAAATTTGGGTGA	93444
Qy	3459	GCCAACTCGCAATTTCTACTACAGCAATTTGAGACCAAGTTAGATTATTTGAAATATTATAGA	3518
Db	93443	GCCAACTCGCAATTTCTACTACAGCAATTTGAGACCAAGTTAGATTATTTGAAATATTATAGA	93384
Qy	3519	GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATGGGATACTTT	3578
Db	93383	GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATAGAACATGGGATACTTT	93324
Qy	3579	AGAACTGAATTTGACAGGSCATATTAGTTGATGAATGAGTCAATTTGAGTCTTATTAATA	3638
Db	93323	AGAACTGAATTTGACAGGSCATATTAGTTGATGAATGAGTCAATTTGAGTCTTATTAATA	93264
Qy	3639	GCCATGTATCATAAATTACCAAGTGAAGCTGGTGGAAACATATATGTTCTCCATTTTACAGTTA	3698
Db	93263	GCCATGTATCATAAATTACCAAGTGAAGCTGGTGGAAACATATATGTTCTCCATTTTACAGTTA	93204
Qy	3699	AGGAATAATAATGACAGATTAATATGTTGTTCTGTGCATGCCACAAATCCCTTTCTAAGGA	3758
Db	93203	AGGAATAATAATGACAGATTAATATGTTGTTCTGTGCATGCCACAAATCCCTTTCTAAGGA	93144
Qy	3759	AGACTGCCCTACTATAGCAGTTTTTATATTTTGTCAATTTATGAATATATGAATGAATGA - GAG	3818
Db	93143	AGACTGCCCTACTATAGCAGTTTTTATATTTTGTCAATTTATGAATATATGAATGAATGA - GAG	93085
Qy	3819	TTCTGGTACCTCTCTCTTTTACAAATATTTGGTGTGTGCCAGTATTTTTTCCCTTTTTTAAC	3878
Db	93084	TTCTGGTACCTCTCTCTTTTACAAATATTTGG-----TGTGTGCAGTATTTTTTCCCTTTTTTA	93029
Qy	3879	CMTTCCCAATTCGGGTGTAGTGGATGTTTCCATTTTGGGTTTTTAAATTTGTATATCCCT	3938
Db	93028	CCATTTCCAATTCGGGTGTAGTGGATG-----TTTCAATTTGGTTTTTAAATTTGTATATCCCT	92974
Qy	3939	GATAGCTATAATTTGGGTCTATAGAAATTTCTTTATACATTTCTAGATCAGAGTCTCTTGYCGG	3998
Db	92973	GATAGCTATAATTTGGGTCTATAGAAATTTCTTTATACATTTCTAGATCAGAGTCTCTTGYCGG	92914
Qy	3999	ATATACGTATTGAGATATTACACCTAGTCTGTGGCTTGACCTGTTTTTCTTTATGCTCTTTTG	4058
Db	92913	ATATATGATTGAGATATTACACCTAGTCTGTGGCTTGACCTGTTTTTCTTTATGCTCTTTTG	92854
Qy	4059	ATGAATAGAAGTTTTTAAATTTTGAACAAGTCAAAATTAATTTTTTCTTTTGTGTAATTT	4118
Db	92853	ATGAATAGAAGTTTTTAAATTTTGAACAAGTCAAAATTAATTTTTTCTTTTGTGTAATTT	92794
Qy	4119	TTTTCTCTCCAAATTTAAACCCCAAGATTTTCAGATATTCTGCTCTATTATATAAACCTTTATA	4178
Db	92793	TTTTCTCTCCAAATTTAAACCCCAAGATTTTCAGATATTCTGCTCTATTATATAAACCTTTATA	92734
Qy	4179	TTTTTATATTTGTGATCTACCTTGAAATTTGATAATGATGTTGTGAAATTTATGATCAGGTT	4238
Db	92733	TTTTTATATTTGTGATCTACCTTGAAATTTGATAATGATGTTGTGAAATTTATGATCAGGTT	92674
Qy	4239	CTTTTTTCCCCCATACAAGTATCCAGTCATTTGTAACACTGTTTATTTGAAGAATTTATCC	4298
Db	92673	CTTTTTTCCCCCATACAAGTATCCAGTCATTTGTAACACTGTTTATTTGAAGAATTTATCC	92614
Qy	4299	TTTCTCTCAATTAATTTACCTTGCCCAATTTAGTAAAAAATCAATTTAAACCATRM	4348
Db	92613	TTTCTCTCAATTAATTTACCTTGCCCAATTTAGTAAAAAATCAATTTAAACCATAA	92564
RESULT 10			
AF049895/c			
LOCUS			
DEFINITION Homo sapiens chromosome 8 multiple clones map q21.3, complete			
sequence.			
ACCESSION AF049895			
VERSION AF049895.3 GI:42600578			
KEYWORDS HTG			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
REFERENCE 1 (bases 1 to 331864)			
AUTHORS Varon, R., Visinga, C., Platzer, M., Cerosaletti, K.M., Chrzanowska, K.H., Saar, K., Beckmann, G., Seemanova, E., Cooper, P.R., Nowak, N.J., Stumm, M., Weemaes, C.M., Gatti, R.A., Wilson, R.K., Digweed, M., Rosenthal, A., Sperling, K., Concannon, P. and Reis, A. Nibrin, a novel DNA double-strand break repair protein, is mutated in Nijmegen breakage syndrome			
TITLE			



[illegible]



misc_feature	2127..2549	/note="similar to EST N66848 (NID:gl218973) za46d10.s1"
repeat_region	3500..3784	/rpt_family="Alu"
repeat_region	3799..4122	/rpt_family="MER1_type"
repeat_region	4114..4180	/rpt_family="MER1_type"
repeat_region	4182..4279	/rpt_family="Alu"
repeat_region	4289..4425	/rpt_family="Alu"
repeat_region	4430..4509	/rpt_family="MER1_type"
repeat_region	4546..5169	/rpt_family="MER1_type"
repeat_region	5190..5343	/rpt_family="L1"
repeat_region	5455..5745	/rpt_family="Alu"
repeat_region	6014..6289	/rpt_family="MER2_type"
repeat_region	6290..6518	/rpt_family="L1"
repeat_region	6534..6824	/rpt_family="Alu"
repeat_region	7558..7871	/rpt_family="Alu"
repeat_region	7938..8238	/rpt_family="Alu"
repeat_region	9286..9467	/rpt_family="MIR"
repeat_region	9941..10012	/rpt_family="Alu"
repeat_region	12133..12441	/rpt_family="Alu"
repeat_region	12513..12804	/rpt_family="Alu"
repeat_region	12808..12912	/rpt_family="L1"
repeat_region	12913..13213	/rpt_family="Alu"
repeat_region	13214..13638	/rpt_family="L1"
repeat_region	13639..13750	/rpt_family="L2"
repeat_region	14231..14529	/rpt_family="Alu"
repeat_region	16063..16506	/rpt_family="L1"
repeat_region	17682..17990	/rpt_family="Alu"
repeat_region	18183..18704	/rpt_family="ERV"
repeat_region	18757..18921	/rpt_family="MIR"
repeat_region	18939..19063	/rpt_family="Alu"
repeat_region	19103..19178	/rpt_family="MIR"
repeat_region	19312..19448	/rpt_family="MIR"
repeat_region	19626..19961	/rpt_family="L2"
repeat_region	20170..20245	/rpt_family="MER1_type"
repeat_region	20787..20974	/rpt_family="L1"
repeat_region	21556..21787	

repeat_region	21793..21917	/rpt_family="L1"
repeat_region	21918..22210	/rpt_family="Alu"
repeat_region	22211..22380	/rpt_family="Alu"
repeat_region	22401..22804	/rpt_family="L1"
repeat_region	22971..23110	/rpt_family="L2"
repeat_region	23617..23881	/rpt_family="Alu"
repeat_region	24264..24633	/rpt_family="MaLR"
repeat_region	24697..24908	/rpt_family="MIR"
repeat_region	25116..25179	/rpt_family="L1"
repeat_region	28103..29649	/rpt_family="L1"
repeat_region	29677..31685	/rpt_family="L1"
repeat_region	31689..31818	

Query Match

42.3%; Score 1861.2; DB 9; Length 107549;

Best Local Similarity 94.1%; Pred. No. 0;

Matches 2023; Conservative 9; Mismatches 27; Indels 91; Gaps 6;

QY	2278	TCCTTTTAGATACAAATCCTTATTTAAAAAGGAGAGATACTAGAGATTTTAAAAAGAG	2337
Db	68939	TCCTTGCAGATACAAATCCTTATTTAAAAAGGAGAGATACTAGAGATTTTAAAAAGAG	68880
QY	2338	CCATGGAAAACTTCCTAGTAGAGCATCTACTTCAGGCCAACAGGTTATATGAATATATA	2397
Db	68879	CCATGGAAAACTTCCTAGTAGAGCATCTACTTCAGGCCAACAGGTTATATGAATATATA	68820
QY	2398	GTGTATAGAAGCGATTTAAAGTTACAATGTTTTATGGCTTAAATTTATTAATAAATGCA	2457
Db	68819	GTGTATAGAAGCGATTTAAAGTTACAATGTTTTATGGCTTAAATTTATTAATAAATGCA	68760
QY	2458	CAAACTTTTGATTCCTTTTGATGTAAACAATTTGTTGTCGTTTTTCAGGCTTTGTCATTG	2517
Db	68759	CAAACTTTTGATTCCTTTTGATGTAAACAATTTGTTGTCGTTTTTCAGGCTTTGTCATTG	68700
QY	2518	CATCTTTTTCATTTTAAATGTGTTTGTATTATAAATAGTTAATAATAGTACAGTTC	2577
Db	68699	CATCTTTTTCATTTTAAATGTGTTTGTATTATAAATAGTTAATAATAGTACAGTTC	68640
QY	2578	AAATTTCTAAATRTACGTAAGGTAAAGGACTAAAGTCCACCTTCACCATTTGTCCTAGCT	2637
Db	68639	AAATTTCTAAATRTACGTAAGGTAAA - GACTAAAGTCCACCTTCACCATTTGTCCTAGCT	68581
QY	2638	ACT-----	2640
Db	68580	ACTTGGTTCCCTCAGAAAAAAATTCATGATACATCTTTTATGAATCTTTTCCAGGGAT	68521
QY	2641	-----TATTTTAAATTAATTCCTACACAAATGATAGCATTAAC	2678
Db	68520	TTTTGAGTCTTATTCAAATTCCTATATTTTAAATAATTTCTTACACAAATGATAGCATTAAC	68461
QY	2679	ATATGCAGTGTCTACACCTTGCTTTTACTTTAGTAGAGTTAAAAATTTATAGGAATATC	2738
Db	68460	ATATGCAGTGTCTACACCTTGCTTTTACTTAGT - AGATTAATAATTTATAGGNATATC	68402
QY	2739	AATATAATGTTTTTAAATAATTTTCTTTTCCATTTATGCTGTAGTCTTACCTAAACTCTGG	2798
Db	68401	AATATAATGTTTTTAAATAATTTTCTTTTCCATTTATGCTGTACTCTTACCTAAACTCTGG	68342
QY	2799	TGATCCAAACAAAATGGCTTCAGTGGTGCAGATGTACCTACATGTTATTTCTAGTACTAG	2858
Db	68341	TGATCCAAACAAAATGGCTTCAGTGGTGCAGATGTACCTACATGTTATTTCTAGTACTAG	68282

Qy	2859	AAACTGAAGACCATGTGGAGACTTCATCAAAACATGGGTTTAGTTTTACCAGAAATGGAAA	2918
Db	68281	AAACTGAAGACCATGTGGAGACTTCATCAAAACATGGGTTTAGTTTTACCAGAAATGGAAA	68222
Qy	2919	GACCTGTACCCCTTTTTTGGTGGTCTTACTAGAGCTGGTGGGTGCTGTGTTTGAGCTTAAT	2978
Db	68221	GACCTGTACCCCTTTTTTGGTGGTCTTACTAGAGCTGGTGGGTGCTGTGTTTGAGCTTAAT	68162
Qy	2979	TAGAGTCCTAGTTTTCTACTTATAAAGTAGAAATGGTGAGATTTGTTTTCTTTTCTACC	3038
Db	68161	TAGAGTCCTAGTTTTCTACTTATAAAGTAGAAATGGTGAGATTTGTTTTCTTTTCTACC	68102
Qy	3039	KTAAGGGGAGATGGTAAGAAACAATGAATGTCTTTTTTCAAACTTTATTTGACAAGTGAT	3098
Db	68101	TTAAAGGGGAGATGGTAAGAAACAATGAATGTCTTTTTTCAAACTTTATTTGACAAGTGAT	68042
Qy	3099	TTCAAGTCTGTGTTTCAAAAATATATTATCATGTACTGTGATCAGCAAGAAAGGAGTCCCA	3158
Db	68041	TTCAAGTCTGTGTTTCAAAAATATATTATCATGTACTGTGATCAGCAAGAAAGGAGTCCCA	67982
Qy	3159	GTCAGAAGTCACTACAACCTGATTAGTTGTTTAGAGAAATGAGAAATGCAACAGTCAGGAA	3218
Db	67981	GTCAGAAGTCACTACAACCTGATTAGTTGTTTAGAGAAATGAGAAATGCAACAGTCAGGAA	67922
Qy	3219	GGAGGCCATATTTCCATGACTTCCCTTGTAAACAGAAAGCAACAGAAAGGCAAGAGGCTG	3278
Db	67921	GGAGGCCATATTTCCATGACTTCCCTTGTAAACAGAAAGCAACAGAAAGGCAAGAGGCTG	67862
Qy	3279	GCCTCTACATCACTCTCACTCCCTCCAAATCTGTGGAAAGTGCACTACTTTGCCAGAACCA	3338
Db	67861	GCCTCTACATCACTCTCACTCCCTCCAAATCTGTGGAAAGTGCACTACTTTGCCAGAACCA	67802
Qy	3339	ATTAACTTACTTCCAAAGTTCTGGCTCTTTCGAGGTGGAACTCCAGCTGCAAGGGAGTTAG	3398
Db	67801	ATTAACTTACTTCCAAAGTTCTGGCTCTTTCGAGGTGGAACTCCAGCTGCAAGGGAGTTAG	67742
Qy	3399	GGAAATGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTAGGGAACAGAAATTTGGGTGA	3458
Db	67741	GGAAATGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTAGGGAACAGAAATTTGGGTGA	67682
Qy	3459	GCCAAATCGCAATTTCTACTACAGGCATTTGAGACCAGTTAGATTATGAATATATAGA	3518
Db	67681	GCCAAATCGCAATTTCTACTACAGGCATTTGAGACCAGTTAGATTATGAATATATAGA	67622
Qy	3519	GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATGAAACATGGGATCTTT	3578
Db	67621	GAGTTATGAACACTTAAATTTATGATAGTGGTATGACATTTGGATGAAACATGGGATCTTT	67562
Qy	3579	AGAAGTAGAATTCAGAGGCATATTAGTTGATGAAATGGAGTCATTTGAGTCTTTTAATA	3638
Db	67561	AGAAGTAGAATTCAGAGGCATATTAGTTGATGAAATGGAGTCATTTGAGTCTCTTAATA	67502
Qy	3639	GCCATGTATCATTAATTTACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTA	3698
Db	67501	GCCATGTATCATTAATTTACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTA	67442
Qy	3699	AGGAATATATGACAGATTTAATATTGTTTCTGTGCATGCCACAAATCCCTTTCTAAGGA	3758
Db	67441	AGGAATATATGACAGATTTAATATTGTTTCTGTGCATGCCACAAATCCCTTTCTAAGGA	67382
Qy	3759	AGACTGCCCTACTATAGCAGTTTTTATATTTCTCAATTTATGAATATATGAATGAGGAG	3818
Db	67381	AGACTGCCCTACTATAGCAGTTTTTATATTTCTCAATTTATGAATATATGAATGAGGAG	67323
Qy	3819	TTCTGTGACTCTCTGCTTTTACAATATTGGGTGTTGTCAGATATTTTCCCTTTTAAAC	3878
Db	67322	TTCTGTGACTCTCTGCTTTTACAATATTGGGTGTTGTCAGATATTTTCCCTTTTAAAC	67267
Qy	3879	CMTTTCCCAATTCGGGTGCTGAGTGGATGTTTCCATTTTGGGTTTTAAATTTGTATATCCCT	3938
Db	67266	CCATTCCAATTCGGGTGCTGAGTGGATGTTTCCATTTTGGGTTTTAAATTTGTATATCCCT	67212
Qy	3939	GATAGCTATAAATTTGGGCTATAGAAAATCTTTTATACATTTAGATGCAAGTCTCTTGYCGG	3998

Db 67 CGCGGTTGACGTCGGCCCGCAGCCCTGAGGAGCCGACCGATGTGGAAACTGCTGCCGCG 126

Qy 73 CGCGGCGCGGCGAGGAGAACCATACAGACTTTTGACTGGCTTGAGTACGTTGTTGG 132

Db 127 CGCGGCGCGGCGAGGAGAACCATACAGACTTTTGACTGGCTTGAGTACGTTGTTGG 186

Qy 133 AAGGAAAACTGTGCCATCTCTAATTCGAAATGATCAGTCGATCAGCCGGAATCATGCTGT 192

Db 187 AAGGAAAACTGTGCCATCTCTAATTCGAAATGATCAGTCGATCAGCCGGAATCATGCTGT 246

Qy 193 GTTAACTGCTAACTTTTCTGTAAACCAACCTGAGTCAAAACAGATGAAATCCCTGTATTGAC 252

Db 247 GTTAACTGCTAACTTTTCTGTAAACCAACCTGAGTCAAAACAGATGAAATCCCTGTATTGAC 306

Qy 253 ATTAAAAAGATAATCTTAAGTATGGTACCTTTGTTAATGAGGAAAAAATCAGAAATGGCTTT 312

Db 307 ATTAAAAAGATAATCTTAAGTATGGTACCTTTGTTAATGAGGAAAAAATCAGAAATGGCTTT 366

Qy 313 TTCCCGAACTTTGAAGTCCGGGGATGATATTACTTTTGGAGTGTCTTGGAAAGTAAATTCAG 372

Db 367 TTCCCGAACTTTGAAGTCCGGGGATGATATTACTTTTGGAGTGTCTTGGAAAGTAAATTCAG 426

Qy 373 AATPAGATGAGCCCTTTGGTTGCATGCTCTCTTGTTTAGATGTCTCTGGAAAACTGC 432

Db 427 AATPAGATGAGCCCTTTGGTTGCATGCTCTCTTGTTTAGATGTCTCTGGAAAACTGC 486

Qy 433 TTTAAATCAAGCTATATTTGCACTTGAGGAATTTACTGTAAACAATTTGGACGAGAATG 492

Db 487 TTTAAATCAAGCTATATTTGCACTTGAGGAATTTACTGTAAACAATTTGGACGAGAATG 546

Qy 493 CACTCACTTGTCTATGCTATCAGTGAAGTTACCATTTAAAAACAATATGTGCATCAATTTG 552

Db 547 CACTCACTTGTCTATGCTATCAGTGAAGTTACCATTTAAAAACAATATGTGCATCAATTTG 606

Qy 553 TGGACGTCCTCAATTTGTAAGCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA 612

Db 607 TGGACGTCCTCAATTTGTAAGCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA 666

Qy 613 GAAGACGCTCCACAAATTTGAAAGTTTACCACCTCTTGATGAACCATCTATTGGAAG 672

Db 667 GAAGACGCTCCACAAATTTGAAAGTTTACCACCTCTTGATGAACCATCTATTGGAAG 726

Qy 673 TAAAAATGTTGATCTGTGAGGCGGAGGAAAGAAACAAATCTTCAAAGGGGAAACATT 732

Db 727 TAAAAATGTTGATCTGTGAGGCGGAGGAAAGAAACAAATCTTCAAAGGGGAAACATT 786

Qy 733 TATATTTTGAATGCCAAACAGCATAAAGAAATTTGAGTTCGCGAGTTGTCTTTGGAGGTGG 792

Db 787 TATATTTTGAATGCCAAACAGCATAAAGAAATTTGAGTTCGCGAGTTGTCTTTGGAGGTGG 846

Qy 793 GGAAGCTAGTTGATTAACAGAGAGAAATGAAGAAACATATTTCTTTTGGCTCCGGG 852

Db 847 GGAAGCTAGTTGATTAACAGAGAGAAATGAAGAAACATATTTCTTTTGGCTCCGGG 906

Qy 853 AAGCTGTGTTGTTGATACAGAAATAACAACTCACAGACCTTAATTTCTTGACTGTGAGAA 912

Db 907 AAGCTGTGTTGTTGATACAGAAATAACAACTCACAGACCTTAATTTCTTGACTGTGAGAA 966

Qy 913 GAAATGGAATTCAGTCAATATGATATGCTTCCAAAGGCAAGGTCTTGTAGACCTATTCTCTGA 972

Db 967 GAAATGGAATTCAGTCAATATGATATGCTTCCAAAGGCAAGGTCTTGTAGACCTATTCTCTGA 1026

Qy 973 AGCAGAAATTTGAAATTCGGCGTGATTTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG 1032

Db 1027 AGCAGAAATTTGAAATTCGGCGTGATTTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG 1086

Qy 1033 CCATCCAGTACAGGATTTAAAGCAACAACCTCAGGACCAAGCTTTTCAAGGGCGTGTG 1092

Db 1087 CCATCCAGTACAGGATTTAAAGCAACAACCTCAGGACCAAGCTTTTCAAGGGCGTGTG 1146

Qy 1093 AGTTGATGAAAACTAATATGCAAGCGCCCGAGTGAACTACTACAACTAGCTAGCTGACAC 1152

Db 1147 AGTTGATGAAAACTAATATGCAAGCGCCCGAGTGAACTACTACAACTAGCTAGCTGACAC 1206

Qy 1153 AGAATCAGAGCAAGCAGATACATGGGATTTGAGTGAAGGCCAAAGAAATCAAAGTCTC 1212

Db 1207 AGAATCAGAGCAAGCAGATACATGGGATTTGAGTGAAGGCCAAAGAAATCAAAGTCTC 1266

Qy 1213 CAAAATGAAACAAAAATTCAGAAATGCTTTTCAAGACGACCCCACTGTTAAAGGAGTCTCTG 1272

Db 1267 CAAAATGGAACAAAA----- 1282

Qy 1273 CAAAACAAGCTCTAATAATAATAGTATGGTATCAAATATCTTTGGCTAAGATGAGAATCCC 1332

Db 1283 ----- 1282

Qy 1333 AAACTATCAGCTTTTCAACCAACTAAATTCGCAAGTATAATAATAAGCTAAAGATAGGCTTTC 1392

Db 1283 -----CACCAACTAAATTCGCAAGTATAATAATAAGCTAAAGATAGGCTTTC 1328

Qy 1393 TCAGCAGCAGCAGACCAACTCCATCAGAAAACTTCTTTTCAGCCGCTTACC-AAAAAAGGG 1451

Db 1329 TCAGCAGCAGCAGACCAACTCCATCAGAAAACTTCTTTTCAGCCGCTTACCANAAGGG 1388

Qy 1452 AAAGGATGAAGAAAAATCAAGAAATGTCTTCATGCAAAATTCAGCAAGAAATAGAAAACTCTT 1511

Db 1389 AAAGGATGAAGAAAAATCAAGAAATGTCTTCATGCAAAATTCAGCAAGAAATAGAAAACTCTT 1448

Qy 1512 GTTCTCTTTTGAACCAACCAACCTGCTACACCTCATTTGCGAAATTAAGGAGCAGC 1571

Db 1449 GTTCTCTTTTGAACCAACCAACCTGCTACACCTCATTTGCGAAATTAAGGAGCAGC 1508

Qy 1572 ATCTATCTGAGAAATGAGCTGTGGACACAAACTTCAGACAAATAACTTTTACAGATACAG 1631

Db 1509 ATCTATCTGAGAAATGAGCTGTGGACACAAACTTCAGACAAATAACTTTTACAGATACAG 1568

Qy 1632 ATTTAAAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGCTGCAGAAAAAGCTAAGAT 1691

Db 1569 ATTTAAAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGCTGCAGAAAAAGCTAAGAT 1628

Qy 1692 CAAATAAAAAAGGGAATGGATGATGTGGCCATAGAGATGAAGTATTTGGAACAGTTAT 1751

Db 1629 CAAATAAAAAAGGGAATGGATGATGTGGCCATAGAGATGAAGTATTTGGAACAGTTAT 1688

Qy 1752 TCAGGACACAAAAACAGAGTTAGAAAATTTGATGTGAAAATTCAAAACAGAGGAGGATG 1811

Db 1689 TCAGGACACAAAAACAGAGTTAGAAAATTTGATGTGAAAATTCAAAACAGAGGAGGATG 1748

Qy 1812 TCAATGTTAGAAAAAGGCCAAGGATGGATATAGAAAAAATAAGCACTTTTCACTGATGAAG 1871

Db 1749 TCAATGTTAGAAAAAGGCCAAGGATGGATATAGAAAAAATAAGCACTTTTCACTGATGAAG 1808

Qy 1872 CAGTACCAAGAAAGTACAAAAATATCTCAAGAAAAATGAAATTTGGGAAGAAACCTGAACTCA 1931

Db 1809 CAGTACCAAGAAAGTACAAAAATATCTCAAGAAAAATGAAATTTGGGAAGAAACCTGAACTCA 1868

Qy 1932 AGGAAGACTCCTATGCTGAGCTAAGAAAAATATCTTAACAATGACAACTTCAGGATGATA 1991

Db 1869 AGGAAGACTCCTATGCTGAGCTAAGAAAAATATCTTAACAATGACAACTTCAGGATGATA 1928

Qy 1992 GTGAGATGCTTCCAAAAAGCTGTTATTGACATGAAATTTTAGTCACTGCTGATTTAAAAACT 2051

Db 1929 GTGAGATGCTTCCAAAAAGCTGTTATTGACTGAAATTTAGTCACTGCTGATTTAAAAACT 1988

Qy 2052 CTACTTCCAGAAATCCGCTCGCATAAATGATGATTATGTTCACTCACTTAAAAAATTTTC 2107

Db 1989 CTACTTCCAGAAATCCGCTCGCATAAATGATGATTATGTTCACTCACTTAAAAAATTTTC 2044

RESULT 13  
AX875862  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS

AX875862 Sequence 10767 from Patent EP1074617.  
AX875862  
AX875862.1 GI:40030598

AX875862 2044 bp DNA linear PAT 17-DEC-2003



QY	1692	CAAAATAAAAAAGCGAAATCGATGATGTGGCCATAGAAGATGAAGTATTGGAACTGTTAT	1751
Db	1629	CAAAATAAAAAAGCGAAATCGATGATGTGGCCATAGAAGATGAAGTATTGGAACTGTTAT	1688
QY	1752	TCAAGGACACAAACACGAGCTTGAAGATTGATGTGAAGTTTCAAAAACAGGAGGAAGATG	1811
Db	1689	TCAAGGACACAAACACGAGCTTGAAGATTGATGTGAAGTTTCAAAAACAGGAGGAAGATG	1748
QY	1812	TCAATGTGAAAAAGGCCAAGCATGATATAGAAACAAATGACACTTTCAGTGATGAAG	1871
Db	1749	TCAATGTGAAAAAGGCCAAGCATGATATAGAAACAAATGACACTTTCAGTGATGAAG	1808
QY	1872	CAGTACCAGAACTAGCAGAAATATCTCAAGAAAATGAAATTTGGGAAGAACTGGAAGTCA	1931
Db	1809	CAGTACCAGAACTAGCAGAAATATCTCAAGAAAATGAAATTTGGGAAGAACTGGAAGTCA	1868
QY	1932	AGGAAGACTCAGTATGTGTGAGCTTAAAGAAATATCTAACATGACAAACTTTCAGGATGATA	1991
Db	1869	AGGAAGACTCAGTATGTGTGAGCTTAAAGAAATATCTAACATGACAAACTTTCAGGATGATA	1928
QY	1992	GTGAGATGCTTCCAAAAAGCTGTTATGCTGAATTTAGATCAGCTGGTGAATTAACAACT	2051
Db	1929	GTGAGATGCTTCCAAAAAGCTGTTATGCTGAATTTAGATCAGCTGGTGAATTAACAACT	1988
QY	2052	CTACTTCCAGAAATCGTCTGGCATAAATGATGATTATGCTCAACTAAAAAATTC	2107
Db	1989	CTACTTCCAGAAATCGTCTGGCATAAATGATGATTATGCTCAACTAAAAAATTC	2044
RESULT 14			
AK001017			
LOCUS	AK001017	2044 bp mRNA linear PRI 30-JAN-2004	
DEFINITION		Homo sapiens cDNA FLJ10155 fis, clone HEMBAL003433, highly similar to Homo sapiens gene for NBS1.	
ACCESSION	AK001017		
VERSION	AK001017.1	GI:7022031	
KEYWORDS		oligo capping; fis (full insert sequence).	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Oabayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H., Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Inose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Teraohima,Y., Suzuki,F., Wakebe,H., Hishigaki,H., Watanabe,T., Goto,Y., Shimizu,F., Waki,Y., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigetani,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.	
TITLE		Complete sequencing and characterization of 21,243 full-length human cDNAs	

JOURNAL	Nat. Genet. 36 (1), 40-45 (2004)		
PUBMED	14702039		
REFERENCES			
AUTHORS	Isogai,T., Ota,T., Hayaishi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sato,H., Sugano,S., Shiratori,A., Sudo,H., Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K., Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.		
TITLE	NEDO human cDNA sequencing project		
JOURNAL			
REFERENCE	3 (bases 1 to 2044)		
AUTHORS	Isogai,T. and Otsuki,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)		
COMMENT	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.		
FEATURES	Location/Qualifiers		
source	1..2044		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="HEMBAL003433"		
	/tissue_type="whole embryo, mainly head"		
	/clone_lib="HEMBAL"		
	/dev_stage="embryo, 10 weeks"		
misc_feature	/note="cloning vector: pME18SFL3"		
	1282..1283		
	/note="118 bases segment is present in AF051334 and AF058696"		
misc_difference	1378		
	/note="compared to AF051334 and AF058696"		
	/replace=""		
ORIGIN			
Query Match	41.7%;	Score 1836.4;	DB 9; Length 2044;
Best Local Similarity	94.3%;	Fred. No. 0;	
Matches 1976;	Conservative 0;	Mismatches 1;	Indels 119; Gaps 2;
QY	13	CGCGGTTGCACTCGGCCCCAGCCCTGAGGAGCCGACCGATGTGGAAACTGCTGCCCGC	72
Db	67	CGCGGTTGCACTCGGCCCCAGCCCTGAGGAGCCGACCGATGTGGAAACTGCTGCCCGC	126
QY	73	CGCGGCCCCGAGGAGAACCATACAGACTTTTGTGCTGGCGTTGAGTAGCTGTGG	132
Db	127	CGCGGCCCCGAGGAGAACCATACAGACTTTTGTGCTGGCGTTGAGTAGCTGTGG	186
QY	133	ANGGAAACTGTGCCATTCTTAATGAAAATGATCAGTCGATCAGCCGAAATCATGCTGT	192
Db	187	AAGGAAACTGTGCCATTCTTAATGAAAATGATCAGTCGATCAGCCGAAATCATGCTGT	246
QY	193	GTTAACTGCTAACTTTTCTGTAACCAACTGAGTCAACAGATGAAATCCCTGTATTGAC	252
Db	247	GTTAACTGCTAACTTTTCTGTAACCAACTGAGTCAACAGATGAAATCCCTGTATTGAC	306
QY	253	ATTAAGAAATATTCATGATGTTGTTGTTTAAATGAGGAAAAATGAGAAATGGCTT	312
Db	307	ATTAAGAAATATTCATGATGTTGTTGTTTAAATGAGGAAAAATGAGAAATGGCTT	366
QY	313	TTCCCGAACTTTGAAGTCGGGGATGGTATTACTTTTCGAGTGTGGGAAGTAAATTCAG	372
Db	367	TTCCCGAACTTTGAAGTCGGGGATGGTATTACTTTTCGAGTGTGGGAAGTAAATTCAG	426
QY	373	AATAGATATGAGCCCTTTGGTTGTCATGCTCTTCTTTGTATTAGATGTCCTCTGGGAAACTGC	432

Db 427 AATAGAGTATGAGCCTTTGGTTGATGCTCTTCTTGTGTTTATAGTGTCTCTGGGAAACTGC 486  
Qy 433 TTTAAATCAAGCTATATTGCAACTTGAGAGTATTTACTGTAAACAAATGGGACAGAAGATG 492  
Db 487 TTTAAATCAAGCTATATTGCAACTTGAGAGTATTTACTGTAAACAAATTTGGACAGAAGATG 546  
Qy 493 CACTCAGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 552  
Db 547 CACTCAGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 606  
Qy 553 TGGACGCTCCAAATTTGTAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTCAGTCCAA 612  
Db 607 TGGACGCTCCAAATTTGTAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTCAGTCCAA 666  
Qy 613 GAAGAGCCTCCCAAAATTTGTAAGCTTATTTTACCACCTCTTTGATGAACCAATCTATTGGAAG 672  
Db 667 GAAGAGCCTCCCAAAATTTGTAAGCTTATTTTACCACCTCTTTGATGAACCAATCTATTGGAAG 726  
Qy 673 TAAAAATGTTGATCTGTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 732  
Db 727 TAAAAATGTTGATCTGTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 786  
Qy 733 TATATTTTTCGAATGCCAAACAGCATTAAGAAATTTGAGTTCCGCAAGTTGTCTTTTGAGGTGG 792  
Db 787 TATATTTTTCGAATGCCAAACAGCATTAAGAAATTTGAGTTCCGCAAGTTGTCTTTTGAGGTGG 846  
Qy 793 GGAAGCTAGGTTGATTAACAGAGAGATGAAGAGAAACATAATTTCTTTTGGCTCCGGG 852  
Db 847 GGAAGCTAGGTTGATTAACAGAGAGATGAAGAGAAACATAATTTCTTTTGGCTCCGGG 906  
Qy 853 AACGTGTGTTGATACAGGAATTAACAACTCACAGACCTTAATTTCTCTGACTGTGAGAA 912  
Db 907 AACGTGTGTTGATACAGGAATTAACAACTCACAGACCTTAATTTCTCTGACTGTGAGAA 966  
Qy 913 GAAATGGAATTCAGTCAATATGATATGCTCCAAAGGCAAGGCTTAGAGCCTATTCTGA 972  
Db 967 GAAATGGAATTCAGTCAATATGATATGCTCCAAAGGCAAGGCTTAGAGCCTATTCTGA 1026  
Qy 973 AGCAGAAATGGAATTTGGCGGTGATTTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG 1032  
Db 1027 AGCAGAAATGGAATTTGGCGGTGATTTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG 1086  
Qy 1033 CCATCCAGTACAGGATTAAGACCAACTCACAGACCAAGCCTTTTCAAGGCGGTGTC 1092  
Db 1087 CCATCCAGTACAGGATTAAGACCAACTCACAGACCAAGCCTTTTCAAGGCGGTGTC 1146  
Qy 1093 AGTTGATGAAAACTAATGCAAGCGCCCGAGTGAACACTACAACTAGCTAGCTGACAC 1152  
Db 1147 AGTTGATGAAAACTAATGCAAGCGCCCGAGTGAACACTACAACTAGCTAGCTGACAC 1206  
Qy 1153 AGAATCAGAGCAAGCAGATACATGGGATTTGAGTGAAGGCGCAAGCCACTGTAAGAGTCTC 1212  
Db 1207 AGAATCAGAGCAAGCAGATACATGGGATTTGAGTGAAGGCGCAAGCCACTGTAAGAGTCTC 1266  
Qy 1213 CAAAATGGAACAAAAATTCAGATGCTTTTCAAGAGCGCAAGCCACTGTAAGAGTCTC 1272  
Db 1267 CAAAATGGAACAAAAATTCAGATGCTTTTCAAGAGCGCAAGCCACTGTAAGAGTCTC 1282  
Qy 1273 CAAAACAGCTCTAATAATAATAGTATGATCAATAACTTTTGGCTAAGATGAGATCCC 1332  
Db 1283 ----- 1282  
Qy 1333 AACTATCAGCTTTTCCAACTAAATTTGCCAAGTATAAATAAAGTAAAGATAGGCGTTC 1392  
Db 1283 -----CACCACCTAAATTTGCCAAGTATAAATAAAGTAAAGATAGGCGTTC 1328  
Qy 1393 TCAGCAGCAGCAGCAACTCCATCAGAACTACTTTTACGCGCTCTACC-AAAAAAGGG 1451  
Db 1329 TCAGCAGCAGCAGCAACTCCATCAGAACTACTTTTACGCGCTCTACCACCAAAAAAAGGG 1388  
Qy 1452 AAGGGATGAGAAATCAGAAATCTCTTCATGCAATCAGCAAGATAGAACTCTT 1511  
Db 1389 AAGGGATGAGAAATCAGAAATCTCTTCATGCAATCAGCAAGATAGAACTCTT 1448

Qy 1512 GTTCTCTTTTAGAACAAAAACAACCTGCTACACCTCAATTTGTGGAAAAAATAAGAGCAGC 1571  
Db 1449 GTTCTCTTTTAGAACAAAAACAACCTGCTACACCTCAATTTGTGGAAAAAATAAGAGCAGC 1508  
Qy 1572 ATCTATCTCAGATGAGCCTGTGGACACAACTCAGACAACTCACTTATTACAGATACAG 1631  
Db 1509 ATCTATCTCAGATGAGCCTGTGGACACAACTCAGACAACTCACTTATTATTACAGATACAG 1568  
Qy 1632 ATTTAAAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGTCGAGAAAAAGCTAAAGAT 1691  
Db 1569 ATTTAAAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGTCGAGAAAAAGCTAAAGAT 1628  
Qy 1692 CAAATAAAAAGGGAAATGGATGATGTGGCCATAGAAGATGAAGTATTGGAAACAGTTAT 1751  
Db 1629 CAAATAAAAAGGGAAATGGATGATGTGGCCATAGAAGATGAAGTATTGGAAACAGTTAT 1688  
Qy 1752 TCAAGGACACAAACACAGAGTTAGAAATTTGATGTGAAAGTTTCAAAACAGGAGGAAGATG 1811  
Db 1689 TCAGGACACAAACACAGAGTTAGAAATTTGATGTGAAAGTTTCAAAACAGGAGGAAGATG 1748  
Qy 1812 TCAATGTTAGAAAAAGGCCAAGGATGGATATAGAAACAAATGACATTTTCAGTGTGAAG 1871  
Db 1749 TCAATGTTAGAAAAAGGCCAAGGATGGATATAGAAACAAATGACATTTTCAGTGTGAAG 1808  
Qy 1872 CAGTACACAGAAAGTAGCAAAATATCTCAAGAAATGAAATTTGGGAAAGAAACGTGAACCTCA 1931  
Db 1809 CAGTACACAGAAAGTAGCAAAATATCTCAAGAAATGAAATTTGGGAAAGAAACGTGAACCTCA 1868  
Qy 1932 AGGAAGACTCATTGTCAGCTTAAGAAATCTTAACAACTTAACAACTTAACAACTTAACAACT 1991  
Db 1869 AGGAAGACTCATTGTCAGCTTAAGAAATCTTAACAACTTAACAACTTAACAACTTAACAACT 1928  
Qy 1992 GTGAGATGCTTCCAAAAAGCTGTTATTGACTGAATTTAGATCAGTGGTATTAAAAAACT 2051  
Db 1929 GTGAGATGCTTCCAAAAAGCTGTTATTGACTGAATTTAGATCAGTGGTATTAAAAAACT 1988  
Qy 2052 CTACTTCCAGAAATCGTCTGGCATAAATGATGATTTATGGTCAACTTAAAAAATTTTC 2107  
Db 1989 CTACTTCCAGAAATCGTCTGGCATAAATGATGATTTATGGTCAACTTAAAAAATTTTC 2044

## RESULT 15

AF218575  
LOCUS Rattus norvegicus Nbs1 (NBS1) mRNA, complete cds.  
DEFINITION AF218575  
ACCESSION AF218575.1 GI:9651647  
VERSION  
KEYWORDS  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
REFERENCE 1 (bases 1 to 2605)  
AUTHORS Lanson,N.A. Jr., Egeland,D.B., Royale,B.A. and Claycomb,W.C.  
TITLE The MRE11-NBS1-RAD50 pathway is perturbed in SV40 large T antigen-immortalized AT-1, AT-2 and HL-1 cardiomyocytes  
JOURNAL Nucleic Acids Res. 28 (15), 2882-2892 (2000)  
MEDLINE 20368653  
PUBMED 10908350  
REFERENCE 2 (bases 1 to 2605)  
AUTHORS Lanson,N.A. Jr., Egeland,D.B., Royale,B.A. and Claycomb,W.C.  
TITLE Direct Submission  
JOURNAL Submitted (22-DEC-1999) Biochemistry and Molecular Biology, Louisiana State University Health Sciences Center, 1901 Perdido St., New Orleans, LA 70112, USA  
FEATURES  
source  
1. .2605  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
1. .2605  
gene

CDS	/gene="NBS1" 40. .2292 /function="required for MRE11, RAD50 and NBS1 colocalization at nuclear foci in response to DNA double-strand breaks" /note="contains forkhead-associated (FHA) and breast cancer carboxy terminal (BRCT) domains; mutated in Nijmegen breakage syndrome" /codon_start=1 /product="Nbs1" /protein_id="AAF91228.1" /db_xref="GI:9651648" /translation="MKWLLPAAASAAPGECRLLAGVEYIVGRKNCALLIENDQISIRN HAYRVNFPVTSLSQDEIPLEITIKONSKYGTFINEKMQNGLSTLTKGDRVTFVGF ESKRVNFPVTSLSQDEIPLEITIKONSKYGTFINEKMQNGLSTLTKGDRVTFVGF TTCALICGRPIVKPESEFLKAVESKTPPEIESFPYDIDEPALGNKVDLSGRER KQIFKGTIVFLNAKHKKLGSVAVFGGGEARLMARGEEEQSPFSAPGTCVVDVIGIT NTQLIITDSORWIKHIMDILQRLHDIRPIPEAIEGLAVIFMTTESYCNPOGQCTPEVK TTTPGSLQGLSANGKVIIPSAIPMNMVTVVADTESPADTCHSLSERPEVKIFGLDQ NSRKLQGCINIKETNSQSNNSNAATLVRGKAPNYQLSPMKPCPAASNNKDWSSQQ QUNSIKNYFQPCSRKREDEENPEQSSCKSRVELSLLEQTQPAEPLSLWKSKOHS QSTELGRASNVSGGDIKPNKSPDSKSFSTEDLARKKEVDLSTEEVLEELLR SYRPELAVQVKEQBADVISIRKPRMDARNOHLNGPVPESNSALORDETEKKDEL OJEAWSKTEVSNLDELQDSRELPRLKLLTFRSLVNVHNSRNLCPLNGRELKNFL KKFKKATCPGAGKPLHIIGGSDLIIGHARKNTELEEWLEWVQKQAKEDSLADDL FRYPNVKRR"											
ORIGIN												
Query Match	32.2%;	Score 1416.4;	DB 10;	Length 2605;								
Best Local Similarity	76.1%;	Pred. No. 3e-243;				Matches 1863;	Conservative 0;	Mismatches 546;	Indels 39;	Gaps 8;		
QY	37	CTGAGAGCGGACCGATGTGGAACTGTGTCGCCGCGCGCGCGCGAGGAGGAACC	96									
Db	24	CTGAGAGCGGACCTCATGTGGAAGCTGCTCCCGCGCGGAAAGTCCGAGGAAACC	83									
QY	97	ATACAGACTTTTGA CTGGCGTTGAGTACGTTGTTGGAAGGAAACATGTGCCATTTCTAAT	156									
Db	84	ATGTCGACTTTTGGCTGGCGTAGAGTACATTTGTTGGGAGGAAACATGTGCCATTTCTAAT	143									
QY	157	TGAAATGATCAGTCGATCAGCCGAAATCATGCTGTGTTAACTGCTAACTTTTCTGTAAC	216									
Db	144	CGAAATGATCAGTCGATCAGTCGAAATCATGCTGTATTTAAGATGAACCTTTCTGTAAC	203									
QY	217	CAACCTGAGTCAACAGATGAATCCCTGTATTGACATTAAGATGAATTTCTAAGTATGG	276									
Db	204	CAGTTTGAGTCAACCGATGAAATTCCTACATTAATAAAGATGAATTTCTAAGTATGG	263									
QY	277	TACCTTTGTTAATGAGGAAATAATGAGATGGCTTTTCCGAACTTTGAGTCGGGGGA	336									
Db	264	AACCTTTATTAACGAAGAAATAATGAGATGGCTTTTCTTCCACTTTTGAAGACAGAGA	323									
QY	337	TGGTATTACTTTTGGAGTGTTCGAAAGTAAATTCAGAAATGAGATGAGCTTTGGTGTGC	396									
Db	324	TAGAGTTACCTTTGGGTGTTCGAAAGTAAATTCAGATGAGATGAGCTTTGGTGTGT	383									
QY	397	ATGCTCTCTTGTGTTAGATGCTCTGGGAAACCTGCTTTAAATCAAGCTATATTGCAACT	456									
Db	384	TTGCTCTCTTGTGTTAGATGCTCTGGGAAACCTGCTTTAAATCAAGCTATATTACAGCT	443									
QY	457	TGGAGGATTTACTGTAAACATTCGACAGAGATGCATCACCTTCTCATGGTATCAGT	516									
Db	444	TGGAGGACTTCACTCGAAACAGCTGGACAGAAAGTGTACTCACCTTGCCATGTCATCAGT	503									
QY	517	GAAAGTTACCATTAACATATATGTGCACCTCATTTGTGGAGCGTCCCAATTTGTAAGCCAGA	576									
Db	504	CAAGTCACCATTAACACTATATGTGCACCTCATTTGTGGAGCTCCCAATTTGTAAGCCAGA	563									
QY	577	ATATTTTACTGAAATCTCTGAAAGAGTTTCACTCCAGAGAGCAGCTCCCAAAATTTGAAG	636									
Db	564	ATATTTTCTGAAATTTCTCAAGAGTTGAGTCCAGACACCACTCCGAAATTTGAAG	623									

QY	637	TTTTTACCACCTCTTGTGATGAACCATATTGTTGAAAGTAAAAATGTTGATCTCTGTCAGGACG	696
DB	624	TTTTTACCACCTCTGATGAACCATATTGTTGAAAGTAAAAATGTTGATCTCTGTCAGGACG	683
QY	697	GCAGGAAAGAAAACAAATCTTCAAGGGGAAAACATTTATATTTTTTGAATGCCAAACAGCA	756
DB	684	ACGTGAAAGAAAACAAATCTTCAAGGGGAAAACATTTGTTGTTTAAATGCCAAGCAGCA	743
QY	757	TAAGAAATTTAGTTCGCGAGTTGTTCTTTGGAGGTGGGAAGCTAGTTGATTAACAGAGA	816
DB	744	TAAGAAATCTCGGCTCGGAGTTGTTCTCGAGGTGGGAAGCAGGCTGATGGCAGAGG	803
QY	817	GAATGAAGAAGAAATAATTTCTTTTGGCTCCGGGAACGTTGTTGTTGATACAGGAAT	876
DB	804	AGTGAAGAGAAACAGAGCTTCTTCTCAGCTCCCGAACCCTGTGTTGTTGAGCTAGGAAT	863
QY	877	AACAAATCTACAGACTTAATTTCTGACTGTGTCAGAGAAATGGATTCAGTCAATTAATGA	936
DB	864	AACAAACACACAGCTCATAATTTACGACTCCAGAGGAAGTGAATTCATTTGATAATGA	923
QY	937	TATGCTCCAAAGGCAAGCTCTTAGACCTATTCTCTGAAGCAGNAATTTGGATTGGCGGTGAT	996
DB	924	TATATCTTCAAGGGCATGATCTCAGACCTATTCTCTGAAGCAGAGATTGGATTGGCTGTTAT	983
QY	997	TTTCACTACTACAAAGAAATTAAGTGTGATCTCTCAGGGCCATCCAGTACAGGATTAAGAC	1056
DB	984	TTTTTATGACTACAGAGATTACTGTAATCCGCGAGGCCAGCTTGTACAGAGTTAAGAC	1043
QY	1057	AACAACTCCAGGACCAAGCTTTTCAAGGCGTGTGATGTTGATGAAAACTAATGCCAAG	1116
DB	1044	GAGGACTCCAGGACCAAGCTTTTCCAGGGCCGTGTCAGCAATGGAAGATCATCCCCAG	1103
QY	1117	CGCCCCAGTGAACACTACAACTAGTGTGTCACAGAGAAATCAGAGCAGGATACATG	1176
DB	1104	TGCGCAATGAATATGACCACTATGCTAGCTGACAGAAATCAGAGCAGGATACATG	1163
QY	1177	---GGATTGAGTGAAGGCCAAAGAAATCAAGTCTCMAAATGGAAACAAAAATTCAG	1233
DB	1164	TATGCTCTGAGTGAAGACCCAGAGAGTAAAGATCTTTGGACTCGATCAAACTCTAG	1223
QY	1234	AATGCTTTCAAGACGCAACCCATGTGTAAGGAGTCTTGCAAAAAAAGCTCTTAATAA	1293
DB	1224	AAAGCTTTTACAAAGGAACATGCAATATAAAGGAGACCTTAATCAAGAGCTCAATTTAA	1283
QY	1294	TAGTATGTTTCAAACTCTTTGGCTTAAGTGAAGTCCCAACTATCAGCTTTCCACAC	1353
DB	1284	CAATCGACGATCAAAATACACTGGTTAGGGGAAAGGCCGAACTATCAGCTTTCCACAA	1343
QY	1354	TAAATTTGCCAAGTATAAATTAAGATGAAGTGAAGGCTTCTCAGCAGCAGCAGCAACTC	1413
DB	1344	GAAATGCTCTGCTGCAAGTAAATAATTAAGATTGGTCTTCTCAGCAGCAGCTG- --AATC	1400
QY	1414	CATCAGAACTACTTTTTCAGCGCTCTACCAAAAAAGGGAAGGATGAAGAAAAATCAAGA	1473
DB	1401	CATCAAAACTACTTTTCCAGCATGCTCCAGAAAAAGGGAAGGATGAAGAAAAATCCAGA	1460
QY	1474	AATGCTTCACTCAATCAGCAGAAATAGAAACGTTGTTGTTCTTTTATAGAACAAACACA	1533
DB	1461	ACAGCTCTCATGCAATCGTCAAGAGTGGAACTGCTTGTGTTCTTTTATAGAACAAACACA	1520
QY	1534	ACCTGCTACACCTCATTTGTGAAAAATAAGGAGCAGCATCTATCTGAGAAATGAGCCTGT	1593
DB	1521	ACCTGAGAACCTCTACTGTGGAAGGACAAAG- --ACCATGAGTCTCAGAGTGAGACTCT	1577
QY	1594	GGACAAACTCAGACAACTATTATTACAGATCAGATTTAAATCTATTGTTGAAAAA	1653
DB	1578	GGCAGGGGATCCAATGCTCATCTGTGGGTGGAATAGACATAAAACCCCAATGGGAAGAG	1637
QY	1654	TTCTGCCAGTAAATCTCATGCTGCAAGAAAGCTAGATCAATAAATAAAGGGAATGGA	1713
DB	1638	TCCTGACCTAGTCAATTTTCCACAGAGATCTTGAAGCAGAAAGAAAGAAAGATG- --	1695
QY	1714	TGATGTGGCCATAGAAGATGAAGTATTGGAACAGTTATTCAAGGACACAAAAACAGAGTT	1773

[illegible]

Search completed: August 30, 2005, 11:06:46  
Job time : 18315 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 04:04:08 ; Search time 2151 Seconds  
(without alignments)

12117.444 Million cell updates/sec

Title: US-09-837-138-1

Perfect score: 4403

Sequence: 1 ttccggcacagcgcggttg.....accgcggtgagctccagct 4403

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 segs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4123.4	93.6	4412	8 ABZ72011	Abz72011 Human nib
2	4105.2	93.2	4386	3 AAZ89047	Aaz89047 Human nib
3	4103.6	93.2	4386	3 AAZ34997	Aaz34997 NBS1 gene
4	1862.8	42.3	65921	3 AAZ89046	Aaz89046 Human nib
5	1836.4	41.7	2044	4 AAH13813	Aah13813 Human cdn
6	671	15.2	752	4 AAH03660	Aah03660 Human cdn
7	564.6	12.8	622	8 ABZ71727	Abz71727 Human can
8	552.2	12.5	646	3 AAA78056	Aaa78056 cdna enco
9	552.2	12.5	646	4 AA128794	Aai28794 Colon tum
10	552.2	12.5	646	8 ABZ32980	Abz32980 Human col
11	544.8	12.4	561	6 ABQ57768	Abq57768 Human col
12	538.6	12.2	544	6 ABL38148	Ab138148 Human col
13	520.4	11.8	543	4 AAH09188	Aah09188 Human cdn
14	507	11.5	587	7 ADS71935	Ads71935 Human kid
15	405.6	9.2	468	9 ACH26622	Ach26622 Human adu
16	402.2	9.1	610	13 ADQ79117	Adq79117 Novel can
17	362	8.2	483	6 ABQ58279	Abq58279 Human col
18	358.4	8.1	396	4 AAF44923	Aaf44923 Human bre
19	239.4	5.4	272	7 ADS70151	Ads70151 Corn seed
20	212.6	4.8	475	6 ABK62896	Abk62896 Rat seque

c	21	212.6	4.8	475	12	ADP72176	Adp72176 Renal tox
	22	193	4.4	195	2	AAT24560	Aat24560 Human gen
	23	167.6	3.8	315	8	ABX43562	Abx43562 Bovine ES
	24	163.4	3.7	6440	4	AAK68763	Aak68763 Human imm
	25	163.4	3.7	6440	4	AAK68764	Aak68764 Human imm
	26	163.4	3.7	6440	4	AAK68762	Aak68762 Human imm
	27	120.2	2.7	121	2	AAT25563	Aat25563 Human gen
	28	119.4	2.7	442	4	AAI38231	Aai38231 Probe #69
	29	119.4	2.7	442	4	AAK32388	Aak32388 Human bon
	30	119.4	2.7	442	4	AAK06689	Aak06689 Human bon
	31	119.4	2.7	442	4	ABS32095	Abs32095 Human liv
	32	119.4	2.7	442	6	ABS07171	Abso7171 Human gen
	33	114	2.6	114	4	AAI51546	Aai51546 Probe #20
	34	114	2.6	114	4	AAK45612	Aak45612 Human bon
	35	114	2.6	114	4	AAK19606	Aak19606 Human bra
	36	114	2.6	114	4	ABS45301	Abs45301 Human liv
	37	114	2.6	114	6	ABS19884	Abs19884 Human gen
	38	97.6	2.2	6237	6	ABL32358	Ab132358 Human imm
	39	92.4	2.1	6237	6	ABL32359	Ab132359 Human imm
	40	87.6	2.0	61739	12	ADQ97343	Adq97343 Human can
	41	81.6	1.9	3285	7	AD831392	Ads81392 Human gen
	42	80	1.8	1300	12	ADP85917	Adp85917 Synthetic
	43	73.6	1.7	539	11	ACN80219	Actn80219 Breast ca
	44	73.6	1.7	228835	12	ADQ97421	Adq97421 Human can
	45	72.2	1.6	7442	4	AA846686	Aas46686 Tumour su

ALIGNMENTS

RESULT 1  
ABZ72011  
ID ABZ72011 standard; cdna; 4412 BP.  
XX  
AC ABZ72011;  
XX  
DT 01-APR-2003 (first entry)  
XX  
DE Human nibrin (NBS) cdna GenBank AF051334.  
XX  
KW Human; cancer; stomach cancer; cytostatic; gene; sa.  
XX  
OS Homo sapiens.  
XX  
PN WO200283899-A1.  
XX  
PD 24-OCT-2002.  
XX  
PF 28-MAR-2002; 2002WO-JF003038.  
XX  
PR 10-APR-2001; 2001JP-00112039.  
PR 21-SEP-2001; 2001JP-00290193.  
XX  
XX (TAKA-) TAKARA BIO INC.  
XX Yoshikawa Y, Okamoto S, Oura T, Mineno J, Asada K, Kato I;  
XX Inoue H, Mori M;  
XX WPI; 2003-093022/08.

Measuring changes in expression of 264 cancer associated genes for detection of stomach cancer and screening of potential anticancer agents.  
Claim 2; Page; 266pp; Japanese.

The invention relates to a method for the detection of cancer in which a change in the expression of 1 or more of 264 specified cancer associated genes, ABZ71694-ABZ71957, or of sequences at least 80% homologous to them in the specimen tissue as compared to normal tissue is observed. The genes are used in detection, diagnosis and treatment of cancer, especially of stomach cancer. The present sequence is that of a cancer associated polynucleotide of the invention. Note: The present sequence was not given in the printed specification but was isolated using the

CC	GenBank accession number given in the DE line									
XX										
SQ	Sequence 4412 BP; 1470 A; 753 C; 853 G; 1336 T; 0 U; 0 Other;									
	Query Match	93.6%;	Score 4123.4;	DB 8;	Length 4412;					
	Best Local Similarity	97.1%;	Pred. No. 0;							
	Matches 4287;	Conservative 7;	Mismatches 28;	Indels 91;	Gaps 6;					
QY	13	CGCGGTGACGTGCGGCCCGCCAGCCCTGAGAGCGGACCGATGTGAAAACCTGTCGCCGC	72							
DB	12	CGCGGTGACGTGCGGCCCGCCAGCCCTGAGAGCGGACCGATGTGAAAACCTGTCGCCGC	71							
QY	73	CGCGGCCCGCGGAGGAGGAGAACCATACAGACTTTTGGCTGGGTGAGTACGTTGCTGG	132							
DB	72	CGCGGCCCGCGGAGGAGGAGAACCATACAGACTTTTGGCTGGGTGAGTACGTTGCTGG	131							
QY	133	AAGGAAAAACCTGTGCCATCTCTAATTGAAAATGATCAGTCGATCAGCGGAAATCATGCTGT	192							
DB	132	AAGGAAAAACCTGTGCCATCTCTAATTGAAAATGATCAGTCGATCAGCGGAAATCATGCTGT	191							
QY	193	GTTAACTGCTAATCTTTCTGTAAACCAACCTGAGTCAAAACAGATGAAATCCCTGTATTGAC	252							
DB	192	GTTAACTGCTAATCTTTCTGTAAACCAACCTGAGTCAAAACAGATGAAATCCCTGTATTGAC	251							
QY	253	ATTAAAGATTAATCTTAAGTATGTTACCTTTGTTAATGAGGAAAAAATGCAAGATGGCTT	312							
DB	252	ATTAAAGATTAATCTTAAGTATGTTACCTTTGTTAATGAGGAAAAAATGCAAGATGGCTT	311							
QY	313	TTCCCGAACTTTGAACTCGGGGATGGTATTACTTTTGGAGTGTGTTGGAAGTAAATTCAG	372							
DB	312	TTCCCGAACTTTGAACTCGGGGATGGTATTACTTTTGGAGTGTGTTGGAAGTAAATTCAG	371							
QY	373	AATAGAGTATGACCTTTGGTTCGATGCTCTCTGTTTAGATGCTCTCGGGAAACTGC	432							
DB	372	AATAGAGTATGACCTTTGGTTCGATGCTCTCTGTTTAGATGCTCTCGGGAAACTGC	431							
QY	433	TTTAAATCAAGCTATATGCACTTGGAGGATTTACTGTAACAATTTGGAACAGAGAAATG	492							
DB	432	TTTAAATCAAGCTATATGCACTTGGAGGATTTACTGTAACAATTTGGAACAGAGAAATG	491							
QY	493	CACCTACCTTTGTCATGGTATCAGTGAAAGTTACCATTTAAAAACAATATGTGCATCATTTG	552							
DB	492	CACCTACCTTTGTCATGGTATCAGTGAAAGTTACCATTTAAAAACAATATGTGCATCATTTG	551							
QY	553	TGGACGTCCAAATGTTAAAGCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA	612							
DB	552	TGGACGTCCAAATGTTAAAGCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA	611							
QY	613	GAGCAGCTCCACAAATTTGAAAGTTTATACCCACCTCTTGATGAACCATCTATTGGAAG	672							
DB	612	GAGCAGCTCCACAAATTTGAAAGTTTATACCCACCTCTTGATGAACCATCTATTGGAAG	671							
QY	673	TAAAAATGTTGATCTGTCAGGACGGCAGGAAAGAAAACAAATCTTCAAAGGGAAAAACATT	732							
DB	672	TAAAAATGTTGATCTGTCAGGACGGCAGGAAAGAAAACAAATCTTCAAAGGGAAAAACATT	731							
QY	733	TATATTTTGAATGCCAAACAGCATGAAGAAATTTGAGTTCCGAGTTGTCTTTGGAGGTGG	792							
DB	732	TATATTTTGAATGCCAAACAGCATGAAGAAATTTGAGTTCCGAGTTGTCTTTGGAGGTGG	791							
QY	793	GGAAGCTAGCTGATTAACAGAGAGAAATGAAGAAGAAACATAATTTCTTTTGGCTCCGGG	852							
DB	792	GGAAGCTAGCTGATTAACAGAGAGAAATGAAGAAGAAACATAATTTCTTTTGGCTCCGGG	851							
QY	853	AACGTGTGTTGTGATACAGGAATAACAAACTCACAGACTTAAATTCCTGACTGTGCAGAA	912							
DB	852	AACGTGTGTTGTGATACAGGAATAACAAACTCACAGACTTAAATTCCTGACTGTGCAGAA	911							
QY	913	GAAATGGATTCAGTCAATTAATGGATATGCTCCAAAGGCAAGGTCTTAGACCTTATCCTGA	972							
DB	912	GAAATGGATTCAGTCAATTAATGGATATGCTCCAAAGGCAAGGTCTTAGACCTTATCCTGA	971							
QY	973	AGCAGAAATTTGGAATTTGGCGGTGATTTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG	1032							

DB	972	AGCAGAAATTTGGATTGGCGGTGATTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGG	1031							
QY	1033	CATCCCCAGTACAGGATTAAGACAACTCCAGGACCAAGCCCTTTTCACAAGGGGTGTC	1092							
DB	1032	CCATCCCGTACAGGATTAAGACAACTCCAGGACCAAGCCCTTTTCACAAGGGGTGTC	1091							
QY	1093	AGTTGATGAAAACCTAATGCCAAGCGCCCGTGAACACTCAACATAGCTAGCTGACAC	1152							
DB	1092	AGTTGATGAAAACCTAATGCCAAGCGCCCGTGAACACTCAACATAGCTAGCTGACAC	1151							
QY	1153	AGAATCAGAGCAAGCAGATACATGCGATTTTGAAGTGAAGGCCCAAGAAATCAAAAGTCTTC	1212							
DB	1152	AGAATCAGAGCAAGCAGATACATGCGATTTTGAAGTGAAGGCCCAAGAAATCAAAAGTCTTC	1211							
QY	1213	CAAAATGGAACAAAATTCAGAAATCTTTTCAAGACGCCCACTGTAAAGGAGTCTCTG	1272							
DB	1212	CAAAATGGAACAAAATTCAGAAATCTTTTCAAGATGCCTTTTCAAGATGCCTTTTAAAGGAGTCTCTG	1271							
QY	1273	CAAAACAGCTCTAATAATAGTATGGTATCAATACTTTTGGCTAAGATGAGAAATCCC	1332							
DB	1272	CAAAACAGCTCTAATAATAGTATGGTATCAATACTTTTGGCTAAGATGAGAAATCCC	1331							
QY	1333	AAAATATCAGCTTTTCAACCACTAAATTTGCCAAGTATAAATAAAGATAGAGGCTTC	1392							
DB	1332	AAAATATCAGCTTTTCAACCACTAAATTTGCCAAGTATAAATAAAGATAGAGGCTTC	1391							
QY	1393	TCAGCAGCAGCAGACCAACTCCATCAGAAAATTAATTTTCAAGGAGTCAACAAAAAGGGA	1452							
DB	1392	TCAGCAGCAGCAGACCAACTCCATCAGAAAATTAATTTTCAAGGAGTCAACAAAAAGGGA	1451							
QY	1453	AAGGATGAAGAAAATCAAGAAAATCTTTTCAAGAAAATCAGCAAGAAATAGAAAACGTTCTTG	1512							
DB	1452	AAGGATGAAGAAAATCAAGAAAATCTTTTCAAGAAAATCAGCAAGAAATAGAAAACGTTCTTG	1511							
QY	1513	TTCTCTTTTAGAAAACCAACCACTGTCTACACCTCATTTGTGGAAAAATAAGGAGCAGCA	1572							
DB	1512	TTCTCTTTTAGAAAACCAACCACTGTCTACACCTCATTTGTGGAAAAATAAGGAGCAGCA	1571							
QY	1573	TCATCTCAGAAATGAGCCTGTGGACACAACTCAGACAAATAATTTTACAGATACAGA	1632							
DB	1572	TCATCTCAGAAATGAGCCTGTGGACACAACTCAGACAAATAATTTTACAGATACAGA	1631							
QY	1633	TTTTAAATCTATTGTGAAAAATTTCTGCAGTAAATCTCATGCTCAGAAAAAGCTAAGATC	1692							
DB	1632	TTTTAAATCTATTGTGAAAAATTTCTGCAGTAAATCTCATGCTCAGAAAAAGCTAAGATC	1691							
QY	1693	AAATAAAAAGGGGAAATGGATGTGGCCCATAGAAATGAAGTATTGGAAACAGTTATT	1752							
DB	1692	AAATAAAAAGGGGAAATGGATGTGGCCCATAGAAATGAAGTATTGGAAACAGTTATT	1751							
QY	1753	CAAGACACAAAACAGAGTTAGAAATTTGATGTGAAAGTTCAAAAACAGGAGGAAAGATGT	1812							
DB	1752	CAAGACACAAAACAGAGTTAGAAATTTGATGTGAAAGTTCAAAAACAGGAGGAAAGATGT	1811							
QY	1813	CAATGTTAGAAAAGGCCCAAGATGATATAGAAAACAAATGACACTTTCAGTGTGAAGC	1872							
DB	1812	CAATGTTAGAAAAGGCCCAAGATGATATAGAAAACAAATGACACTTTCAGTGTGAAGC	1871							
QY	1873	AGTACACAGAAAGTAGCAAAAATCTCTCAAGAAAATGAAATTTGGGAAAGAAAACGTGAAC	1932							
DB	1872	AGTACACAGAAAGTAGCAAAAATCTCTCAAGAAAATGAAATTTGGGAAAGAAAACGTGAAC	1931							
QY	1933	GGAAGACTCATATGCTGAGCTTAAAGAAAATATCTAACTAAGTGAACAACTTCAGAGTATAG	1992							
DB	1932	GGAAGACTCATATGCTGAGCTTAAAGAAAATATCTAACTAAGTGAACAACTTCAGAGTATAG	1991							
QY	1993	TCAGATGCTTCCAAAAAGCTGTTTACTGAAATTTAGATCACTGGTGTATTTAAAACTC	2052							
DB	1992	TCAGATGCTTCCAAAAAGCTGTTTACTGAAATTTAGATCACTGGTGTATTTAAAACTC	2051							
QY	2053	TACTTCCAGAAATCCGCTCTGGCATAAATGATGATTTATGGTCAACTTAAAAAATTTCAAGAA	2112							

Db 2052 TACTTCAGAAAATCCATCTCGCATAAATGAATGATTTATGGTCAACTAAAAATTTCAAGAA 2111  
Qy 2113 ATTCAAAAAGGTACATATCTCTGAGCAGGAAATCTCCACACATCATTTGGAGGATCAGA 2172  
Db 2112 ATTCAAAAAGGTACATATCTCTGAGCAGGAAATCTCCACACATCATTTGGAGGATCAGA 2171  
Qy 2173 TCTAAATAGCTCATCTGCTCGAAAGAAATACAGAACTAGAGAGTGGCTAAGGCAGGAAAT 2232  
Db 2172 TCTAAATAGCTCATCTGCTCGAAAGAAATACAGAACTAGAGAGTGGCTAAGGCAGGAAAT 2231  
Qy 2233 GGAGGTACAAAAATCAACATCCAAAGAAAGAGTCTCTTGCTGATGATCTTTTATAGATACAA 2292  
Db 2232 GGAGGTACAAAAATCAACATCCAAAGAAAGAGTCTCTTGCTGATGATCTTTTATAGATACAA 2291  
Qy 2293 TCCTTATTTAAAAAGGAGAAAGATAAATCTGAGGATTTTAAAAAGAGCCATCGAAAACTTC 2352  
Db 2292 TCCTTATTTAAAAAGGAGAAAGATAAATCTGAGGATTTTAAAAAGAGCCATCGAAAACTTC 2351  
Qy 2353 CTAGTAAGCATCTACTTCAGGCCAAACAAAGGTTATATGAATATATAGTGTATAGAGCGAT 2412  
Db 2352 CTAGTAAGCATCTACTTCAGGCCAAACAAAGGTTATATGAATATATAGTGTATAGAGCGAT 2411  
Qy 2413 TTAAGTTACAAATGTTTATGGCCTTAAATTTATTAATAAATGCAACAACTTTGATCT 2472  
Db 2412 TTAAGTTACAAATGTTTATGGCCTTAAATTTATTAATAAATGCAACAACTTTGATCT 2471  
Qy 2473 TTTGATGTAAACAATGTTTGTCTGTTTTCAGGCTTTGTCATTTGTCATCTTTTTCATT 2532  
Db 2472 TTTGATGTAAACAATGTTTGTCTGTTTTCAGGCTTTGTCATTTGTCATCTTTTTCATT 2531  
Qy 2533 TTTAAATGCTTTTGTATTAAATAGTTAAATAGTCACAGTTTCAAAATTTCTAAATRTA 2592  
Db 2532 TTTAAATGCTTTTGTATTAAATAGTTAAATAGTCACAGTTTCAAAATTTCTAAATRTA 2591  
Qy 2593 CGTAAGGTAAGACATAAGTCAACCTTCCACCAATGTCCTAGCTACT----- 2640  
Db 2592 CGTAAGGTAAGACATAAGTCAACCTTCCACCAATGTCCTAGCTACT----- 2650  
Qy 2641 ----- 2640  
Db 2651 GAAAAAAATCATGATACTCATTTCTTATGAATCTTTCCAGGGATTTTTCAGTCCTATTTC 2710  
Qy 2641 -----TATTTTAAATAATTTCTTACACAAATAGATAGCATACATATGCAAGTTCTA 2693  
Db 2711 AAATTCCTATTTTAAATAATTTCTTACACAAATAGATAGCATACATATGCAAGTTCTA 2770  
Qy 2694 CACCTTGCTTTTACTAGTAAGATTAATAATATAGGAATATCAATATAATGTTTTTA 2753  
Db 2771 CACCTTGCTTTTACTTAGT-AGATTAATAATATAGGAATATCAATATAATGTTTTTA 2829  
Qy 2754 ATATTTTCTTTTCCATTAATGCTAGTCTTACCTAAACTCTGGTGATCCAAACAAAAT 2813  
Db 2830 ATATTTTCTTTTCCATTAATGCTAGTCTTACCTAAACTCTGGTGATCCAAACAAAAT 2889  
Qy 2814 GGCTTCAGTGGTGCAGATGTCACCTACATGTTATTTCTAGTACTAGAACTGAAGACCATG 2873  
Db 2890 GGCTTCAGTGGTGCAGATGTCACCTACATGTTATTTCTAGTACTAGAACTGAAGACCATG 2949  
Qy 2874 TGAGACTTCATCAACATGGGTTTATGTTTCCAGGAATGGAAGACCTGTACCCCTTT 2933  
Db 2950 TGAGACTTCATCAACATGGGTTTATGTTTCCAGGAATGGAAGACCTGTACCCCTTT 3009  
Qy 2934 TTGGTGGTCTTACTGAGCTGGGTGCTGTTTGTGAGCTTATTTAGAGTCTCTAGTTTT 2993  
Db 3010 TTGGTGGTCTTACTGAGCTGGGTGCTGTTTGTGAGCTTATTTAGAGTCTCTAGTTTT 3069  
Qy 2994 CCTACTTATAAAGTAGAAATGGTGAGATTTGTTTTCTTTTCTACCTTAAAGGGAGATGGT 3053  
Db 3070 CCTACTTATAAAGTAGAAATGGTGAGATTTGTTTTCTTTTCTACCTTAAAGGGAGATGGT 3129  
Qy 3054 AAGAAACAATGAATGCTTTTTTCAAACTTTATGACAAAGTATTTTCAAGTCTGTGTTT 3113  
Db 3130 AAGAAACAATGAATGCTTTTTTCAAACTTTATGACAAAGTATTTTCAAGTCTGTGTTT 3189

Qy 3114 AAAAAATATATTTCAATGTACCTGTGATCCAGCAAGAGGAGTTCCAGTCAAGAGTCACTAC 3173  
Db 3190 AAAAAATATATTTCAATGTACCTGTGATCCAGCAAGAGGAGTTCCAGTCAAGAGTCACTAC 3249  
Qy 3174 AACTGATTAGTTGTTTAGAGAAATGGAACAGTGAAGGAATGGAGCCATATTTTCC 3233  
Db 3250 AACTGATTAGTTGTTTAGAGAAATGGAACAGTGAAGGAATGGAGCCATATTTTCC 3309  
Qy 3234 ATGACTTCTCTTCTGATAACAGAGCAACAGAGGCAAGAGGCTGCGCTCTACATCACTC 3293  
Db 3310 ATGACTTCTCTTCTGATAACAGAGCAACAGAGGCAAGAGGCTGCGCTCTACATCACTC 3369  
Qy 3294 TCACCTTCCAAATCTTGTGGAGTGCATCTACTTGCAGAACCAAAATTAACCTTCTTCCA 3353  
Db 3370 TCACCTTCCAAATCTTGTGGAGTGCATCTACTTGCAGAACCAAAATTAACCTTCTTCCA 3429  
Qy 3354 AGTTCTGGCTGCTTGCAGGTGGAACTCCAGCTGCAAGGGAGTTAGGGAATGAAGTCTT 3413  
Db 3430 AGTTCTGGCTGCTTGCAGGTGGAACTCCAGCTGCAAGGGAGTTAGGGAATGAAGTCTT 3489  
Qy 3414 TTTTTAAAAAGCTTCTCAGACCTTCTAGGGAACAGAAATTTGGGTGAGCCAATCTGCAATTT 3473  
Db 3490 TTTTTAAAAAGCTTCTCAGACCTTCTAGGGAACAGAAATTTGGGTGAGCCAATCTGCAATTT 3549  
Qy 3474 CTACTACAGGCATTTGAGACCCAGTTAGATTATTGAAATATTTATAGAGAGTTTATGAACACTT 3533  
Db 3550 CTACTACAGGCATTTGAGACCCAGTTAGATTATTGAAATATTTATAGAGAGTTTATGAACACTT 3609  
Qy 3534 AAATTAATGATAGTGGTATGACATTTGGATAGAACATGGGATATCTTTAGAGTAGAATGAC 3593  
Db 3610 AAATTAATGATAGTGGTATGACATTTGGATAGAACATGGGATATCTTTAGAGTAGAATGAC 3669  
Qy 3594 AGGGCATATTTAGTTGATGAATGGAGTCAATTTGAGTCTTTTATAGCCATGTATCATAAT 3653  
Db 3670 AGGGCATATTTAGTTGATGAATGGAGTCAATTTGAGTCTCTTTATAGCCATGTATCATAAT 3729  
Qy 3654 TACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATAATGGAC 3713  
Db 3730 TACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATAATGGAC 3789  
Qy 3714 AGATTAATATTTGTTTCTGTATGCCCAAACTCCCTTTCTAAGGAAGAGCTGCCCTACTAT 3773  
Db 3790 AGATTAATATTTGTTTCTGTATGCCCAAACTCCCTTTCTAAGGAAGAGCTGCCCTACTAT 3849  
Qy 3774 AGCAGTTTTTATATTTGTCAATTTTATGAATATAATGAATGA- GAGTTCTGGTACCTCCTG 3833  
Db 3850 AGCAGTTTTTATATTTGTCAATTTTATGAATATAATGAATGA- GAGTTCTGGTACCTCCTG 3908  
Qy 3834 TCTTTACAAATATTTGGGTGTTGTCCAGTATTTTCCCTTTTAAACCMWTTCCCAATTCGGG 3893  
Db 3909 TCTTTACAAATATTTGG- - - - -TGTGTGTCAGTATTTTCCCTTTTAAACCAATCCCAATCGGTG 3964  
Qy 3894 TGTGTAAGTGGATGTTTCCATTTGGGTTTTTAATTTGTATATATCCCTGATAGCTATAATG 3953  
Db 3965 TGTAGTGAAGT- - - - -TTTCAATTTGGTTTTTAATTTGTATATATCCCTGATAGCTATAATG 4019  
Qy 3954 GTCATAGAAATTTCTTTATACATCTAGATGCAAGTCTCTTGGGATATACGTTATTTGAGA 4013  
Db 4020 GTCATAGAAATTTCTTTATACATCTAGATGCAAGTCTCTTGGGATATATGTTATTTGAGA 4079  
Qy 4014 TATTACACCTAGTCTGTGGCTTGACTGTTTCTTTATGTCTTTTGTATGAATAGAAGTTTT 4073  
Db 4080 TATTACACCTAGTCTGTGGCTTGACTGTTTCTTTATGTCTTTTGTATGAATAGAAGTTTT 4139  
Qy 4074 AAATTTTGAAGGTCAAAATTTATTTTCTTTTGTGTTTGTGATATTTTCTCTCCAAATTT 4133  
Db 4140 AAATTTTGAAGGTCAAAATTTATTTTCTTTTGTGATATTTTCTCTCTCCAAATTT 4199  
Qy 4134 AACCCTCAAGATTTGAGATTTCTGCTCTATTTATATAAACTTTATATATTTTATATTTTGTGA 4193  
Db 4200 AACCCTCAAGATTTGAGATTTCTGCTCTATTTATATAAACTTTATATATTTTATATTTTGTGA 4259

QY 4194 TCTACCTTGAATTGATATGTTGTAATTATGATCAGGGTCTCTTTTCCCCCAT 4253  
 DB |||||  
 DB 4260 TCTACCTTGAATTGATATGTTGTAATTATGATCAGGGTCTCTTTTCCCCCAT 4319  
 QY 4254 ACAAGTATCCAGTCATTGTAACACTGTTTATTGAAAGAAATTATCTTCTCATTAAT 4313  
 DB |||||  
 DB 4320 ACAAGTATCCAGTCATTGTAACACTGTTTATTGAAAGAAATTATCTTCTCATTAAT 4379  
 QY 4314 ACCTTCCCAATTAGTAAAAAATCAATTAACCAT 4346  
 DB |||||  
 DB 4380 ACCTTCCCAATTAGTAAAAAATCAATTAACCAT 4412

## RESULT 2

AAZ89047

ID AAZ89047 standard; DNA; 4386 BP.

AC

XX

AAZ89047;

XX 01-JUN-2000 (first entry)

XX Human nibrin DNA.

DE

XX

KW Nibrin; human; DNA double strand break repair protein; diagnosis;  
 therapy; Nijmegen Breakage Syndrome; gene therapy; ds.

XX

OS

XX

XX

FH

FT

FT

FT

XX

PN

XX

XX

PD

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

QY 148 CATTCTAATTTGAAATGATCAGTCGATCAGCGGAAATCATGCTGTGTTAACTGCTAACTT 207  
 DB |||||  
 DB 121 CATTCTGATTTGAAATGATCAGTCGATCAGCGGAAATCATGCTGTGTTAACTGCTAACTT 180  
 QY 208 TTTCTGTACCAACCTGAGTCACAAAGATGAAATCCCTGTATTGACATTTAAAGATAATTC 267  
 DB |||||  
 DB 181 TTTCTGTACCAACCTGAGTCACAAAGATGAAATCCCTGTATTGACATTTAAAGATAATTC 240  
 QY 268 TAAGTATGGTACCTTTGTTAATGAGGAAAAATGAGAAATGCGCTTTTCCGAACTTTGAA 327  
 DB |||||  
 DB 241 TAAGTATGGTACCTTTGTTAATGAGGAAAAATGAGAAATGCGCTTTTCCGAACTTTGAA 300  
 QY 328 GTCGGGGATGGTATTACTTTTGGAGTGTGTTGGAGTAAATTCAGAAATGAGTATGAGCC 387  
 DB |||||  
 DB 301 GTCGGGGATGGTATTACTTTTGGAGTGTGTTGGAGTAAATTCAGAAATGAGTATGAGCC 360  
 QY 388 TTTGGTTGCATGCTCTTCTGTTTAGATGTCCTGCGGAAAACTGCTTTAAATCAAGCTAT 447  
 DB |||||  
 DB 361 TTTGGTTGCATGCTCTTCTGTTTAGATGTCCTGCGGAAAACTGCTTTAAATCAAGCTAT 420  
 QY 448 ATTGCAACTTTGGAGGATTTACTGTAAACAATTTGGACAGAAGATGCACCTCACCTTGTGAT 507  
 DB |||||  
 DB 421 ATTGCAACTTTGGAGGATTTACTGTAAACAATTTGGACAGAAGATGCACCTCACCTTGTGAT 480  
 QY 508 GGTATCAGTGAAGTTTACCATTTAAACAATATGTCACCTCATTTTGTGACGTCCTCAATTTGT 567  
 DB |||||  
 DB 481 GGTATCAGTGAAGTTTACCATTTAAACAATATGTCACCTCATTTTGTGACGTCCTCAATTTGT 540  
 QY 568 AAAGCCAGATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAAGAGAGCCTCCACA 627  
 DB |||||  
 DB 541 AAAGCCAGATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAAGAGAGCCTCCACA 600  
 QY 628 AATTGAAAGTTTACCCACCTCTTGATGAACCATCTATTGGAAGTAAATTTGATGCT 687  
 DB |||||  
 DB 601 AATTGAAAGTTTACCCACCTCTTGATGAACCATCTATTGGAAGTAAATTTGATGCT 660  
 QY 688 GTCAGGACGCGAGGAAAGAAACAAATCTTCAAAGGGAAAAACATTTATTTTGAATGC 747  
 DB |||||  
 DB 661 GTCAGGACGCGAGGAAAGAAACAAATCTTCAAAGGGAAAAACATTTATTTTGAATGC 720  
 QY 748 CAAACAGCATTAAGAAATTTAGTTCGCGAGTTGCTTTTGGAGTGGGGAAGCTAGGTTGAT 807  
 DB |||||  
 DB 721 CAAACAGCATTAAGAAATTTAGTTCGCGAGTTGCTTTTGGAGTGGGGAAGCTAGGTTGAT 780  
 QY 808 AACGAGACAGATGAAGAAAGAACATAATTTCTTTTGGCTCCGGGAAACGTTGTTGA 867  
 DB |||||  
 DB 781 AACGAGACAGATGAAGAAAGAACATAATTTCTTTTGGCTCCGGGAAACGTTGTTGA 840  
 QY 868 TACAGGAATAACAACTCACAGACCTTAAATTCCTGACTGTGAGAAAGAAATGGATTCAGTC 927  
 DB |||||  
 DB 841 TACAGGAATAACAACTCACAGACCTTAAATTCCTGACTGTGAGAAAGAAATGGATTCAGTC 900  
 QY 928 AATAATGATATGCTCCAAAGGCAAGGTCCTTAGACCTATTCTGAGCAGAAAAATTTGGATT 987  
 DB |||||  
 DB 901 AATAATGATATGCTCCAAAGGCAAGGTCCTTAGACCTATTCTGTAAGCAGAAAAATTTGGATT 960  
 QY 988 GCGGTGATTTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGSCCATCCAGTACAGG 1047  
 DB |||||  
 DB 961 GCGGTGATTTTTCATGACTACAAAGAAATTTACTGTGATCCTCAGGSCCATCCAGTACAGG 1020  
 QY 1048 ATTAAGACAACTCCAGGACCAAGCCTTTTCAAGCGGTGTCAGTTGATGAAAAACT 1107  
 DB |||||  
 DB 1021 ATTAAGACAACTCCAGGACCAAGCCTTTTCAAGCGGTGTCAGTTGATGAAAAACT 1080  
 QY 1108 AATGCAAGCGCCAGTGAACTATCAACATACCTAGCTGACACAGAAATCAGACCAAGC 1167  
 DB |||||  
 DB 1081 AATGCAAGCGCCAGTGAACTATCAACATACCTAGCTGACACAGAAATCAGACCAAGC 1140  
 QY 1168 AGATACATGGGATTTTGTGTAAGGCGCCAAAGAAATCAAGTCTCCAAAAATGGAACAAA 1227  
 DB |||||  
 DB 1141 AGATACATGGGATTTTGTGTAAGGCGCCAAAGAAATCAAGTCTCCAAAAATGGAACAAA 1200  
 QY 1228 ATTCAGAAATGCTTTTCAAGAGCGCACCCCACTGTAAAGGAGTCTCTGCAAAAAACAAGCTCTAA 1287

SQ Sequence 4386 BP; 1467 A; 744 C; 845 G; 1330 T; 0 U; 0 Other;

Query Match

Best Local Similarity 93.2%; Score 4105.2; DB 3; Length 4386;

Matches 4270; Conservative 7; Mismatches 30; Indels 91; Gaps 6;

QY 28 GCCCAGCCCTGAGGAGCGGACCGATGTGGAACCTGCTGCCGCGCGCGGCCCGGCGAGG 87

DB 1 GCCCAGCCCTGAGGAGCGGACCGATGTGGAACCTGCTGCCGCGCGCGGCCCGGCGAGG 60

QY 88 AGGAGAACCATACAGACTTTTGTGCTGCGGTGAGTACGTTGTTGGAAGAAAACTGTGC 147

DB 61 AGGAGAACCATACAGACTTTTGTGCTGCGGTGAGTACGTTGTTGGAAGAAAACTGTGC 120



```
QY 3369 CAGGTGGAACCTCAGCTGCAAGGAGCTTAGGGAATGAAGGTCTTTTTTAAAGCTTCT 3428
Db 3419 CAGGTGGAACCTCAGCTGCAAGGAGCTTAGGGAATGAAGGTCTTTTTTAAAGCTTCT 3478
QY 3429 CAGCCCTTCCTAGGGAACAGAAAATGGGTGAGCCAAATCTGCAATTTCTACTACAGGCATTG 3488
Db 3479 CAGCCCTTCCTAGGGAACAGAAAATGGGTGAGCCAAATCTGCAATTTCTACTACAGGCATTG 3538
QY 3489 AGACCAGTTAGATTATGAAATATATATAGAGAGTTATGAACACTTAAATATATGATAGTGG 3548
Db 3539 AGACCAGTTAGATTATGAAATATATATAGAGAGTTATGAACACTTAAATATATGATAGTGG 3598
QY 3549 TATGACATTTGGATAGAACATGGGATCTTTAGAGTAGAATTCACAGGCGCATATTAGTTG 3608
Db 3599 TATGACATTTGGATAGAACATGGGATCTTTAGAGTAGAATTCACAGGCGCATATTAGTTG 3658
QY 3609 ATGAAATGGAGTCATTTGAGTCTTTTAATAGCCATCTATCAATTAACCAAGTGAAGCTG 3668
Db 3659 ATGAAATGGAGTCATTTGAGTCTTTTAATAGCCATCTATCAATTAACCAAGTGAAGCTG 3718
QY 3669 GTGGAACATATGCTCTCCATTTTACAGTTAAGGAATATTAATGGACAGATTAATATTGTTT 3728
Db 3719 GTGGAACATATGCTCTCCATTTTACAGTTAAGGAATATTAATGGACAGATTAATATTGTTT 3778
QY 3729 TCTGTCTATGCCCAACATCCCTTCTTAAGGAGTCCCTACTATAGCAGTCTTTTATATT 3788
Db 3779 TCTGTCTATGCCCAACATCCCTTCTTAAGGAGTCCCTACTATAGCAGTCTTTTATATT 3838
QY 3789 TGTCAATTTATGAATATAATGAATGAGGAGTTCTGCTACTCTCTGTTTACAAATATTG 3848
Db 3839 TGTCAATTTATGAATATAATGAATGAGGAGTTCTGCTACTCTCTGTTTACAAATATTG 3897
QY 3849 GGTGTTGTCCAGTATTTTCCCTTTTAAACCMFTCCCAATTCGGGTGTAGTGAATGT 3908
Db 3898 G-----TGTTGTAGTATTTTCTTTTAAACCATTCGAATCGGTGTAGTGAATG----- 3948
QY 3909 TTCCATTTGGTGTATTTTGTATATCCCTGATAGCTATTAATTTGGTGTATAGAAATCTT 3968
Db 3949 TTTCAATTTGGTGTATTTTGTATATCCCTGATAGCTATTAATTTGGTGTATAGAAATCTT 4008
QY 3969 TATACATTTAGATGCAAGTCTCTGCGGATATACGTATTGAGATATTACACCTAGTCT 4028
Db 4009 TATACATTTAGATGCAAGTCTCTGCGGATATATGATTTAGATATTACACCTAGTCT 4068
QY 4029 GTGGCTTGACTGTTTCTTTATAGTCTTTTGAATAGAGTTTAAATTTTGAAGAAGT 4088
Db 4069 GTGGCTTGACTGTTTCTTTATGCTTTTGAATAGAGTTTAAATTTTGAAGAAGT 4128
QY 4089 CAAATTTATTTTCTTTTGTGATATTTTCTCTCCAAATTAACCCCAAGATTCTCA 4148
Db 4129 CAAATTTATTTTCTTTTGTGATATTTTCTCTCCAAATTAACCCCAAGATTCTCA 4188
QY 4149 GATATTTCTGCTCTATTATATAAACTTTATATTTTATATTTTGTGATCTACCTTGAATTGA 4208
Db 4189 GATATTTCTGCTCTATTATATAAACTTTATATTTTATATTTTGTGATCTACCTTGAATTGA 4248
QY 4209 TATGATTTGTGAATATAGATTCAGGGTCTTTTTTCCCCCATTAAGATATCCAGTCA 4268
Db 4249 TATGATTTGTGAATATAGATTCAGGGTCTTTTTTCCCCCATTAAGATATCCAGTCA 4308
QY 4269 TTGTACACTGTTTATTTGAAGAAATATCTTTCTTCATTAATTAATTAATTAATTAATTAAT 4328
Db 4309 TTGTACACTGTTTATTTGAAGAAATATCTTTCTTCATTAATTAATTAATTAATTAATTAAT 4368
QY 4329 AAAAAATCAATTAACCAT 4346
Db 4369 AAAAAATCAATTAACCAT 4386
```

RESULT 3  
AAZ34997  
ID AAZ34997 standard; cDNA; 4386 BP.

```
XX AAZ34997;
XX 28-FEB-2000 (first entry)
XX NBS1 gene associated with Nijmegen breakage syndrome.
XX NBS1 gene; nibrin; Nijmegen breakage syndrome; diagnosis; human;
XX gene therapy; cancer; microcephaly; mental retardation;
XX primary ovarian failure; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 26..2290
XX /*tag= a
XX polyA_signal 2420..2425
XX /*tag= b
XX repeat_region 3853..4386
XX /*tag= c
XX /note= "LMC/D genome wide repeat"
XX WO9955716-A1.
XX 04-NOV-1999.
XX 27-APR-1999; 99WO-US009036.
XX 27-APR-1999; 98US-0083269P.
XX (VIRG-) VIRGINIA MASON RES CENT.
XX Concannon PJ, Vissinga CS, Cerosaletti KM, Varon R, Sperling K;
XX Reis A;
XX WPI; 2000-062015/05.
XX P-PSDB; AAY32373.
XX Novel gene useful for detecting mutations or polymorphisms, and
XX diagnosing certain pathological conditions in Nijmegen Breakage syndrome
XX patients.
XX Claim 2; Fig 2A-C; 58pp; English.
XX This is the nucleotide sequence of cDNA for the NBS1 gene that is
XX associated with the Nijmegen breakage syndrome (NBS). It includes a
XX coding region for a 754-amino acid protein, nibrin (see AAY32373). The
XX gene maps to a 1 cm region on chromosome 8q21. The invention is based on
XX the discovery that the gene contains mutations in all NBS patients. These
XX mutations include deletions and insertions that result in frameshift, as
XX well as point mutations. Specific mutations associated with the NBS
XX phenotype include 657del5, 698del4, 835del4, 1142insT, 1142delC, 976C>T,
XX 681delT and 900del25. Polymorphisms include 553 G/C, 1197 T/C, 2016 A/G,
XX 102 G/A, IVS 5+9 T/C, IVS5+51delT, IVS9+18 C/T and IVS-7A/G. It is an
XX object of the invention to detect a mutation or polymorphism in NBS
XX patients, and hence to diagnose a predisposition to a pathological
XX condition such as cancer, microcephaly, mental retardation, and primary
XX ovarian failure, based on detection of a mutation in the NBS1 gene. It is
XX also an object of the invention to treat NBS by replacing the mutated
XX gene in a NBS patient by gene therapy. Recombinant vectors, genetically
XX engineered host cells, a method for producing nibrin polypeptide, an
XX antibody that specifically binds to the polypeptide, and a method for
XX diagnosing NBS are claimed. Primers (see AAZ34998-5035) used for
XX detection a mutation in NBS1 are also claimed
XX Sequence 4386 BP; 1467 A; 747 C; 841 G; 1331 T; 0 U; 0 Other;
XX
```

Query Match 93.2%; Score 4103.6; DB 3; Length 4386;  
Best Local Similarity 97.1%; Pred No. 0;  
Matches 4269; Conservative 7; Mismatches 31; Indels 91; Gaps 6;

QY 28 GCCCCAGCCCTGAGAGCGGACCGACCATGTGGAATCTGTGCCGCCGCGGCGCCGAGG 87  
|||||

Db 1 GCCCCAGCCCTGAGGAGCCGGAACCGATGTGGAAACTGTCTGCCCGCCGCGCCGAGG 60  
Qy 88 AGGAGAACCATACAGACTCTTTGACTGGCGTTGAGTAGCTGTGTGGAGGAGAAAACCTGTGC 147  
Db 61 AGGAGAACCATACAGACTCTTTGACTGGCGTTGAGTAGCTGTGTGGAGGAGAAAACCTGTGC 120  
Qy 148 CATCTAAATTGAAAATGATCAGTCGATCAGCCGGAATCATGCTGTGTAACTCTCAACTT 207  
Db 121 CATCTGATTGAAAATGATCAGTCGATCAGCCGGAATCATGCTGTGTAACTCTCAACTT 180  
Qy 208 TTCTGTAAACCAACCTGAGTCAACACAGATGAATCCCTGTATTGCACATTAAGAGATAATTC 267  
Db 181 TTCTGTAAACCAACCTGAGTCAACACAGATGAATCCCTGTATTGCACATTAAGAGATAATTC 240  
Qy 268 TAAGTAGGTGATCTTTGTTAATGAGGAAAAATGCAAGATGGCTTTTCCCGAACTTTGAA 327  
Db 241 TAAGTAGGTGATCTTTGTTAATGAGGAAAAATGCAAGATGGCTTTTCCCGAACTTTGAA 300  
Qy 328 GTCGGGGGATGGTATTACTCTTTGGAGTGTGTTGGAAGTAATTCAGAAATGAGTATGAGCC 387  
Db 301 GTCGGGGGATGGTATTACTCTTTGGAGTGTGTTGGAAGTAATTCAGAAATGAGTATGAGCC 360  
Qy 388 TTTGGTGTGATGCTCTCTGTTTGTAGATGCTCTGGGAAAACTGCTTTAAATCAAGCTAT 447  
Db 361 TTTGGTGTGATGCTCTCTGTTTGTAGATGCTCTGGGAAAACTGCTTTAAATCAAGCTAT 420  
Qy 448 ATTGCAACTTTGGAGGATTTACTGTAAACAAATTGGACAGAGAAATGCACCTTGTCTAT 507  
Db 421 ATTGCAACTTTGGAGGATTTACTGTAAACAAATTGGACAGAGAAATGCACCTTGTCTAT 480  
Qy 508 GGTATCAGTGAAGTTACCAATTAACAAATATGTGCACCTCATTTTGGAGCGTCCAAATGT 567  
Db 481 GGTATCAGTGAAGTTACCAATTAACAAATATGTGCACCTCATTTTGGAGCGTCCAAATGT 540  
Qy 568 AAGCCAGATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAGAGCAGCCCTCCACA 627  
Db 541 AAGCCAGATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAGAGCAGCCCTCCACA 600  
Qy 628 AATTGAAAGTTTATCCACCTCTTGATGAACCATCTATTGGAAGTAAAAATGTTGATCT 687  
Db 601 AATTGAAAGTTTATCCACCTCTTGATGAACCATCTATTGGAAGTAAAAATGTTGATCT 660  
Qy 688 GTCAGACGCGAGGAAGAAAAACAATCTTCAAAGGGAACAAATTTATTTTGAATGC 747  
Db 661 GTCAGACGCGAGGAAGAAAAACAATCTTCAAAGGGAACAAATTTATATTTTGAATGC 720  
Qy 748 CAACAGCATAGAAATAGATTCCGAGTTGCTTTGGAGGTGGGAGAGCTAGGTTGAT 807  
Db 721 CAACAGCATAGAAATAGATTCCGAGTTGCTTTGGAGGTGGGAGAGCTAGGTTGAT 780  
Qy 808 AACAGAGAGAAATGAAGAGAACATAATTTCTTTTGGCTCCGGGAACGTGTGTTGTCGA 867  
Db 781 AACAGAGAGAAATGAAGAGAACATAATTTCTTTTGGCTCCGGGAACGTGTGTTGTCGA 840  
Qy 868 TACAGGAATAACAAACTCAAGACCTTAAATTCCTGATCTGTCAAGAGAAATGGATTGAGTC 927  
Db 841 TACAGGAATAACAAACTCAAGACCTTAAATTCCTGATCTGTCAAGAGAAATGGATTGAGTC 900  
Qy 928 AATAATGGATATGCTCCAAAGGCAAGCTCTAGACCTTATTCCTGAAGCAGAAATTCGATT 987  
Db 901 AATAATGGATATGCTCCAAAGGCAAGCTCTAGACCTTATTCCTGAAGCAGAAATTCGATT 960  
Qy 988 GGGGGTATTTTTCATGACTACAAAGAAATTTACTGTGATCTCTCAGGGCCATCCAGTACAGG 1047  
Db 961 GGGGGTATTTTTCATGACTACAAAGAAATTTACTGTGATCTCTCAGGGCCATCCAGTACAGG 1020  
Qy 1048 ATTAAGACAACAACCTCCAGGACCAAGCCCTTTCAAGGCGTGTCTGATTGATGAAAAACT 1107  
Db 1021 ATTAAGACAACAACCTCCAGGACCAAGCCCTTTCAAGGCGTGTCTGATTGATGAAAAACT 1080  
Qy 1108 AATGCCAAGGCCCGGAGTGAACACTACACATACGTAGCTGACACAGAAATCAGAGCAAGC 1167  
Db 1081 AATGCCAAGGCCCGGAGTGAACACTACACATACGTAGCTGACACAGAAATCAGAGCAAGC 1140

Qy 1168 AGATACATGGGATTTGAGTGAAAGGCCAAAGAAATCAAAGTCTCCAAATGGAACAAA 1227  
Db 1141 AGATACATGGGATTTGAGTGAAAGGCCAAAGAAATCAAAGTCTCCAAATGGAACAAA 1200  
Qy 1228 ATTCAAGATGCTTTCAAGACGACCCCACTGCTAAAGAGTCTCTGCAAAAACAGCTCTAA 1287  
Db 1201 ATTCAAGATGCTTTCAAGATGACCCCACTGCTAAAGAGTCTCTGCAAAAACAGCTCTAA 1260  
Qy 1288 TAATAATAGTATGTTATCAATACATTTTGGCTAAGATGAGAAATCCCAAACTATCAGCTTTC 1347  
Db 1261 TAATAATAGTATGTTATCAATACATTTTGGCTAAGATGAGAAATCCCAAACTATCAGCTTTC 1320  
Qy 1348 ACCAACTAAATTTGCCAAGTATAAATAAAGTAAAGATAGGGCTTTCTCAGCAGCAGCAGAC 1407  
Db 1321 ACCAACTAAATTTGCCAAGTATAAATAAAGTAAAGATAGGGCTTTCTCAGCAGCAGCAGAC 1380  
Qy 1408 CAACTCCATCAGAAACTACTTTTCAGCCGTCTACCAAAAAAGGAAAGGATGAAGAAA 1467  
Db 1381 CAACTCCATCAGAAACTACTTTTCAGCCGTCTACCAAAAAAGGAAAGGATGAAGAAA 1440  
Qy 1468 TCAAGAAATGCTCTTCATGCAAAATCAGCAAGAAATAGAAACGTCTTGTCTCTTTTGAACA 1527  
Db 1441 TCAAGAAATGCTCTTCATGCAAAATCAGCAAGAAATAGAAACGTCTTGTCTCTTTTGAACA 1500  
Qy 1528 AACACAACCTGCTACACCTCTATTGTGGAAAAATTAAGGACGACATCTATCTGAGAAATGA 1587  
Db 1501 AACACAACCTGCTACACCTCTATTGTGGAAAAATTAAGGACGACATCTATCTGAGAAATGA 1560  
Qy 1588 GCCTGTGGACAAAACCTCAGACAATACTTATTATTACAGATACAGATTTAAAAATCTATTGT 1647  
Db 1561 GCCTGTGGACAAAACCTCAGACAATACTTATTATTACAGATACAGATTTAAAAATCTATTGT 1620  
Qy 1648 GAAATTTCTGCCAGTAAATCTCATGCTGCAGAAAGCTAAAGTCAAATTAAGAAAGGGA 1707  
Db 1621 GAAATTTCTGCCAGTAAATCTCATGCTGCAGAAAGCTAAAGTCAAATTAAGAAAGGGA 1680  
Qy 1708 AATGATGATGTGGCCATAGAAAGATGAAGTATTGGAAACAGTATTATCAAGGACACAAAAAC 1767  
Db 1681 AATGATGATGTGGCCATAGAAAGATGAAGTATTGGAAACAGTATTATCAAGGACACAAAAAC 1740  
Qy 1768 AGAGTTAGAAATTTGATGTGAAAGTTCAAAAACAGGAGGAAGATGTCAATTTAGAAAAAG 1827  
Db 1741 AGAGTTAGAAATTTGATGTGAAAGTTCAAAAACAGGAGGAAGATGTCAATTTAGAAAAAG 1800  
Qy 1828 GCCAAGGATGGATATAGAAACAAATGACACTTTTCAGTGTATGAAGCAGTACCAAGAAAGTAG 1887  
Db 1801 GCCAAGGATGGATATAGAAACAAATGACACTTTTCAGTGTATGAAGCAGTACCAAGAAAGTAG 1860  
Qy 1888 CAAAATATCTCAAGAAAAATGAAATTTGGGAAGAAAACGTGAACCTCAAGGAAGACTCACTATG 1947  
Db 1861 CAAAATATCTCAAGAAAAATGAAATTTGGGAAGAAAACGTGAACCTCAAGGAAGACTCACTATG 1920  
Qy 1948 GTCAGCTAAAGAAATATCTTAAACAATGACAAAACCTCAGGAATGATAGTGAATGCTTCCAAA 2007  
Db 1921 GTCAGCTAAAGAAATATCTTAAACAATGACAAAACCTCAGGAATGATAGTGAATGCTTCCAAA 1980  
Qy 2008 AAGAGCTGTTATGACTGAAATTTTAGATCACTGGTGTATTAATAACTCTACTTCCAGAAATCC 2067  
Db 1981 AAGAGCTGTTATGACTGAAATTTTAGATCACTGGTGTATTAATAACTCTACTTCCAGAAATCC 2040  
Qy 2068 GTCTGGCATAAATGATGATTTATGGTCACTAAAAAATTTCAAAGAAATTTCAAAGAGTCACTAC 2127  
Db 2041 ATCTGGCATAAATGATGATTTATGGTCACTAAAAAATTTCAAAGAAATTTCAAAGAGTCACTAC 2100  
Qy 2128 ATATCTGGAGCAGGAAAACTTCCACACATCAATTCAGGAGTCAAGTCAATAGCTCATCA 2187  
Db 2101 ATATCTGGAGCAGGAAAACTTCCACACATCAATTCAGGAGTCAAGTCAATAGCTCATCA 2160  
Qy 2188 TCCTCGAAGATACAGAACTAGAGAGTGGCTTAAGGCAAGGAAATGGAGGTACAAATCA 2247  
Db 2161 TCCTCGAAGATACAGAACTAGAGAGTGGCTTAAGGCAAGGAAATGGAGGTACAAATCA 2220

QY 2248 ACATGCAAAAGAGAGTCTCTGCTGATGATCTTTTATAGATACAATCCTTATTTAAABAG 2307  
DB |||||  
DB 2221 ACATGCAAAAGAGAGTCTCTGCTGATGATCTTTTATAGATACAATCCTTATTTAAABAG 2280  
QY 2308 GAGAAGATAACTGAGGATTTTAAAGAGAGCCATGGAAAACCTTCTAGTAGACATCTTAC 2367  
DB |||||  
DB 2281 GAGAAGATAACTGAGGATTTTAAAGAGAGCCATGGAAAACCTTCTAGTAGACATCTAC 2340  
QY 2368 TTGAGGCCAACAGGTTATATGAATATATAGTGTATAGAGCGATTTAAGTTACAATGTT 2427  
DB |||||  
DB 2341 TTGAGGCCAACAGGTTATATGAATATATAGTGTATAGAGCGATTTAAGTTACAATGTT 2400  
QY 2428 TTATGCGCTAAATTTATTAATAAATGACACAAAATTTGATCTTTTCTGATGTACAAAT 2487  
DB |||||  
DB 2401 TTATGCGCTAAATTTATTAATAAATGACACAAAATTTGATCTTTTCTGATGTACAAAT 2460  
QY 2488 TGTGTTGTCGTTTTGAGGCTTTGTCATTCATCTCTTTTTTCAATTTTAAATGCTGTTTTG 2547  
DB |||||  
DB 2461 TGTGTTGTCGTTTTGAGGCTTTGTCATTCGATCTTTTTCATTTTAAATGCTGTTTTG 2520  
QY 2548 TTTTATTAATAGTTAATATAGTACACAGTTCAAAAATTTCTAAATATAGTAAAGGAC 2607  
DB |||||  
DB 2521 TTTTATTAATAGTTAATATAGTACACAGTTCAAAAATTTCTAAATATAGTAAAGGTA 2579  
QY 2608 TAAAGTCACCTTCCACCATGTCCTAGCTACT- 2640  
DB |||||  
DB 2580 TAAAGTCACCTTCCACCATGTCCTAGCTACTGTTGTTCCCTCAGAAAAAATTCATGA 2639  
QY 2641 -----TATTTTTA 2648  
DB |||||  
DB 2640 TACTCATTTCTTATGAATCTTTCCAGGATTTTGTAGTCTATTCAAAATTCCTATTTTTA 2699  
QY 2649 AATAATTTCTTACAAAATGATAGATATAATGAGTGTCTACACCTGCTGTTTTTA 2708  
DB |||||  
DB 2700 AATAATTTCTTACAAAATGATAGATATAATGAGTGTCTACACCTGCTTTTTTA 2759  
QY 2709 CTTTGTAGTAAATTAATAATATAGGAATATCAATATAATGTTTTTAAATTTTCTTTTC 2768  
DB |||||  
DB 2760 CTTAGT-AGATTAATAATTAATAGGAATATCAATATAATGTTTTTAAATTTTCTTTTC 2818  
QY 2769 CATTTAGCTGTAGTCTTACCTAAACTCTGCTGATCCAAAACAAAATGGCTTCAGTGTGCA 2828  
DB |||||  
DB 2819 CATTTAGCTGTAGTCTTACCTAAACTCTGCTGATCCAAAACAAAATGGCTTCAGTGTGCA 2878  
QY 2829 GATGTACCTACATGTTATTCAGTACTAGAAATCGAAGACATGTTGAGACTTCATCAA 2888  
DB |||||  
DB 2879 GATGTACCTACATGTTATTCAGTACTAGAACTGAAGACCATGTGAGACTTCATCAA 2938  
QY 2889 ACATGGGTTTGTGTTTCCACAGAAATGAAAGACCTGTACCCTTTTGTGTTGTTACTG 2948  
DB |||||  
DB 2939 ACATGGGTTTGTGTTTCCACAGAAATGAAAGACCTGTACCCTTTTGTGTTGTTACTG 2998  
QY 2949 AGCTGGGTGGGTGCTCTGTTTTGAGCTTATTTAGAGTCTTAGTCTTCTACTTATAAAGTA 3008  
DB |||||  
DB 2999 AGCTGGGTGGGTGCTCTGTTTTGAGCTTATTTAGAGTCTTAGTCTTCTACTTATAAAGTA 3058  
QY 3009 GAAATGGTGAGATGTTTCTTTTCTACCKTAAAGGAGATGGTAAAGAAAACATGAATG 3068  
DB |||||  
DB 3059 GAAATGGTGAGATGTTTCTTTTCTACCTTAAAGGAGATGGTAAAGAAAACATGAATG 3118  
QY 3069 TCTTTTTTCAAACTTTATTTGACAAGTATTTTCAAGTCTGTGTTCAAAAATATATTCATG 3128  
DB |||||  
DB 3119 TCTTTTTTCAAACTTTATTTGACAAGTATTTTCAAGTCTGTGTTCAAAAATATATTCATG 3178  
QY 3129 TACCTGTGATCCAGCAAGGAGTTCCAGTCAAGAGTCACTPACAACTGATTTAGTTGTT 3188  
DB |||||  
DB 3179 TACCTGTGATCCAGCAAGGAGTTCCAGTCAAGAGTCACTPACAACTGATTTAGTTGTT 3238  
QY 3189 TAGAGAAATGAGAAATGGAACAGTGAAGAAATGGAGGCCATATTTCCATGATCTCCCTTGTA 3248  
DB |||||  
DB 3239 TAGAGAAATGAGAAATGGAACAGTGAAGAAATGGAGGCCATATTTCCATGATCTCCCTTGTA 3298  
QY 3249 AACAGAAGCAACAGAAGGACAGAGGCTGGCCTCTACATCACTCTCACCTTCCAAATCT 3308

DB |||||  
QY 3299 AACAGAGCAACAGAAAGGACAGAGGCTGGCCTCTACATCACTCTCACCTTCCAAATCT 3358  
QY 3309 TGTGAAAGTGCACTCTACTTGCAGAAACCAAAATTAATCTTACTTCCAAAGTTCTGGCTGCTTG 3368  
DB |||||  
DB 3359 TGTGAAAGTGCACTCTACTTGCAGAAACCAAAATTAATCTTACTTCCAAAGTTCTGGCTGCTTG 3418  
QY 3369 CAGGTGGAACTCCAGCTGCNAAGGAGTTAGGAAATGAAGTCTTTTTTAAAGCTTCT 3428  
DB |||||  
DB 3419 CAGGTGGAACTCCAGCTGCNAAGGAGTTAGGAAATGAAGTCTTTTTTAAAGCTTCT 3478  
QY 3429 CAGCCTTCTTAGGAAACAGAAATTTGGGTGAGCCAATCTGCAATTTCTACTACAGGCATTTG 3488  
DB |||||  
DB 3479 CAGCCTTCTTAGGAAACAGAAATTTGGGTGAGCCAATCTGCAATTTCTACTACAGGCATTTG 3538  
QY 3489 AGACAGTTAGATTTAATGAATATATATAGAGTTTATGAACACTTTAAATATATGATAGTG 3548  
DB |||||  
DB 3539 AGACAGTTAGATTTAATGAATATATATAGAGTTTATGAACACTTTAAATATATGATAGTG 3598  
QY 3549 TATGACATTTGGATAGAACATGGGATACCTTTAGNAGTAGAATTTGACAGGGCATATTAGTTG 3608  
DB |||||  
DB 3599 TATGACATTTGGATAGAACATGGGATACCTTTAGNAGTAGAATTTGACAGGGCATATTAGTTG 3658  
QY 3609 ATGAAATGGAGTCACTTTGAGTCTTTTAAATAGCCATGTATCAATAATCAAGTGAAGCTG 3668  
DB |||||  
DB 3659 ATGAAATGGAGTCACTTTGAGTCTTTTAAATAGCCATGTATCAATAATCAAGTGAAGCTG 3718  
QY 3669 GTGGAACATATATGCTTCCATTTTACAGTTAAGGAATATAATGAGCAGATTAATATGTTT 3728  
DB |||||  
DB 3719 GTGGAACATATATGCTTCCATTTTACAGTTAAGGAATATAATGAGCAGATTAATATGTTT 3778  
QY 3729 TCTGTGTCATGCCACCAATCCCTTTCTAAGGAAGACTGCCCTACTATAGCAGTTTTTATATT 3788  
DB |||||  
DB 3779 TCTGTGTCATGCCACCAATCCCTTTCTAAGGAAGACTGCCCTACTATAGCAGTTTTTATATT 3838  
QY 3789 TGTCAATTTATGAATATAATGAATGAGGAGTTCTGGTACCTCTCTCTTTTACAAATATTG 3848  
DB |||||  
DB 3839 TGTCAATTTATGAATATAATGAATGAGGAGTTCTGGTACCTCTCTCTTTTACAAATATTG 3897  
QY 3849 GGTGTTGTCAGTATTTTCCCTTTTAAACCMTTTCCAAATTCGGGTGTGAGTGGATGT 3908  
DB |||||  
DB 3898 G-----TGTGTCAGTATTTTCTTTTAAACCAATTCGGGTGTGAGTGAATG- 3948  
QY 3909 TTCCATTTTGGGTTTTTAAATTTGATATCCCTGATAGCTATAATGCGTCAATAGAAATCTTT 3968  
DB |||||  
DB 3949 TTTCAATTTGGTTTTTAAATTTGATATCCCTGATAGCTATAATTTGGGTCAATAGAAATCTT 4008  
QY 3969 TATACATTTCTAGATCAAGTCTCTTGTGCGGATATACGTATTTGAGATATTTACACCTAGTCT 4028  
DB |||||  
DB 4009 TATACATTTCTAGATCAAGTCTCTTGTGCGGATATATGATTTGAGATATTTACACCTAGTCT 4068  
QY 4029 GTGGCTGACCTGTTTTTCTTTATGCTTTTGTAGTAAATGAGGTTTTTAAATTTTGAACAAGT 4088  
DB |||||  
DB 4069 GTGGCTGACCTGTTTTTCTTTATGCTTTTGTAGTAAATGAGGTTTTTAAATTTTGAACAAGT 4128  
QY 4089 CAAATTTATTTTCTTTTCTTTGATATTTTCTCTCCAATTTTAAACCCCAAGATTTTCA 4148  
DB |||||  
DB 4129 CAAATTTATTTTCTTTTGTGATATTTTTCTCTCCAATTTTAAACCCCAAGATTTTCA 4188  
QY 4149 GATATTTCTGCTCTATTATATAAACTTTTATATTTTATATTTTGTGATCTACCTTTGAATTTGA 4208  
DB |||||  
DB 4189 GATATTTCTGCTCTATTATATAAACTTTTATATTTTATATTTTGTGATCTACCTTTGAATTTGA 4248  
QY 4209 TATGATGTTGTGAATATGGAATCAGGGTTCTTTTTTCCCCCATACAGATATCCAGTCA 4268  
DB |||||  
DB 4249 TATGATGTTGTGAATATGGAATCAGGGTTCTTTTTTCCCCCATACAGATATCCAGTCA 4308  
QY 4269 TTGTAACTGTTTTTATGAAAGATTTATCTCTTCTCAATTAATTTACCTTTGCAATTTAGT 4328  
DB |||||  
DB 4309 TTGTAACTGTTTTTATGAAAGATTTATCTCTTCTCAITTAATTTACCTTTGCCAATTTAGT 4368  
QY 4329 AAAAAATCAATTTAAACCAT 4346  
DB |||||

Db	4369	AAAAATCAATTAAACCAT	4386	
		RESULT 4		
		AAZ89046		
ID	AAZ89046	standard; DNA; 65921 BP.		
XX	AC	AAZ89046;		
XX	AC			
XX	01-JUN-2000	(first entry)		
XX	DT			
XX	DE	Human nibrin DNA.		
XX	Nibrin; human; DNA double strand break repair protein; diagnosis;			
KW	therapy; Nijmegen Breakage Syndrome; gene therapy; ds.			
XX	XX			
OS	Homo sapiens.			
XX	XX			
PN	DE19818680-Cl.			
XX	XX			
PD	09-MAR-2000.			
XX	XX			
PF	27-APR-1998;	98DE-01018680.		
XX	XX			
PR	27-APR-1998;	98DE-01018680.		
XX	XX			
PA	(UYBE ) UNIV BERLIN HUMBOLDT.			
XX	XX			
XX	WPI; 2000-196117/18.			
XX	XX			
PT	A DNA double strand break repair protein, Nibrin, and related DNA useful			
PT	for diagnosis and therapy of Nijmegen Breakage Syndrome and other			
PT	diseases influenced by DNA-double-strand break repair.			
XX	XX			
PS	Claim 4; Fig 1; 32pp; German.			
XX	XX			
CC	This invention describes a novel DNA double strand break repair protein,			
CC	Nibrin. Nibrin and DNA encoding it are useful for diagnosis and/or			
CC	therapy of diseases influenced by repair of DNA-double strand breaks, in			
CC	particular Nijmegen Breakage Syndrome. The product of the invention has			
CC	applications in gene therapy. This sequence encodes the nibrin protein			
CC	described in the invention			
XX	XX			
SQ	Sequence 65921 BP; 19501 A; 11699 C; 12463 G; 22149 T; 0 U; 109 Other;			
	Query Match	42.3%;	Score 1862.8;	DB 3; Length 65921;
	Best Local Similarity	94.1%;	Pred. No. 0;	
	Matches 2024; Conservative	9;	Mismatches 26;	Indels 91; Gaps 6;
QY	2278	TCCTTTTAGATACAATCCCTTATTTAAAGGAGAGAGATAACTGAGGATTTTAAAAAGAG	2337	
Db	58191	TCCTTCAGATACAATCCCTTATTTAAAGGAGAGAGATAACTGAGGATTTTAAAAAGAG	58250	
QY	2338	CCATGAAAAAATTCCTAGTAGCATCTACTTCAGGCCAACAAAGGTTATATGAATATATA	2397	
Db	58251	CCATGAAAAAATTCCTAGTAGCATCTACTTCAGGCCAACAAAGGTTATATGAATATATA	58310	
QY	2398	GTGTATAGAGCGATTTAGTTACATGTTTATGGCCTAAATTTTAAATTAATGAATGCA	2457	
Db	58311	GTGTATAGAGCGATTTAGTTACATGTTTATGGCCTAAATTTTAAATTAATGAATGCA	58370	
QY	2458	CAAAACCTTTGATCTTTTCTGTATGAACAATTTGTTGCTGTTTTCAGGCTTTGTCAATTG	2517	
Db	58371	CAAAACCTTTGATCTTTTCTGTATGAACAATTTGTTGCTGTTTTCAGGCTTTGTCAATTG	58430	
QY	2518	CATCTTTTTCATTTTAAATGTGTTTGTATTAATGAATAGTTTAAATAGTACAGTTC	2577	
Db	58431	CATCTTTTTCATTTTAAATGTGTTTGTATTAATGAATAGTTTAAATAGTACAGTTC	58490	
QY	2578	AAATTTCTAAATRTAGCTAAGGTAAAGGCTAAAGTCAACCTTCCACCATTTGTCCTAGCT	2637	
Db	58491	AAATTTCTAAATGTACGTAAAGGTAAA -GACTAAAGTCAACCTTCCACCATTTGTCCTAGCT	58549	

Db 59629 GCGATGATCATTAATACCAAGTGAAGCTGGTGAACATATGGTCTCCATTTTACAGTTA 59688  
QY 3699 AGGAATATAATGACAGATTAATATATTTCTGTCATGCCCAAAATCCCTTCTTAAGGA 3758  
Db 59689 AGGAATATAATGACAGATTAATATATTTCTGTCATGCCCAAAATCCCTTCTTAAGGA 59748  
QY 3759 AGACTGCCCTACTATAGCAGTTTATATTTCTGTCATTTATGAATATGAATGAGGAG 3818  
Db 59749 AGACTGCCCTACTATAGCAGTTTATATTTCTGTCATTTATGAATATGAATGAG-GAG 59807  
QY 3819 TTCTGCTACTCTCTGCTCTTACAAATATTTGGGTGTTGTCAGATTTTTCCTTTTAAAC 3878  
Db 59808 TTCTGCTACTCTCTGCTCTTACAAATATTTGG- - -TGTGTGTCAGATTTTTCCTTTTAA 59863  
QY 3879 CMTTCCCAATTCGGGTGTAGTGGATGTTTCCATTTTGGGTTTTAAATTTGATATATCCCT 3938  
Db 59864 CCATTCCTCAATTCGGGTGTAGTGGATGTTTCCATTTTGGGTTTTAAATTTGATATATCCCT 59918  
QY 3939 GATAGCTAATTTGGGTGTAGTGGATGTTTCCATTTTGGGTTTTAAATTTGATATATCCCT 3998  
Db 59919 GATAGCTAATTTGGGTGTAGTGGATGTTTCCATTTTGGGTTTTAAATTTGATATATCCCT 59978  
QY 3999 ATATAGCTAATTTGGGTGTAGTGGATGTTTCCATTTTGGGTTTTAAATTTGATATATCCCT 4058  
Db 59979 ATATAGCTAATTTGGGTGTAGTGGATGTTTCCATTTTGGGTTTTAAATTTGATATATCCCT 60038  
QY 4059 ATGAATAGAGTTTAAATTTTGAAGTCAAAATTTTATTTTCTTTTGGTGGATATTT 4118  
Db 60039 ATGAATAGAGTTTAAATTTTGAAGTCAAAATTTTATTTTCTTTTGGTGGATATTT 60098  
QY 4119 TTTTCTCTCAATTTTAAATTTTGAAGTCAAAATTTTATTTTCTTTTGGTGGATATTT 4178  
Db 60099 TTTTCTCTCAATTTTAAATTTTGAAGTCAAAATTTTATTTTCTTTTGGTGGATATTT 60158  
QY 4179 TTTTCTCTCAATTTTAAATTTTGAAGTCAAAATTTTATTTTCTTTTGGTGGATATTT 4238  
Db 60159 TTTTCTCTCAATTTTAAATTTTGAAGTCAAAATTTTATTTTCTTTTGGTGGATATTT 60218  
QY 4239 CTTTCTCTCAATTTTAAATTTTGAAGTCAAAATTTTATTTTCTTTTGGTGGATATTT 4298  
Db 60219 CTTTCTCTCAATTTTAAATTTTGAAGTCAAAATTTTATTTTCTTTTGGTGGATATTT 60278  
QY 4299 TTTTCTCTCAATTTTAAATTTTGAAGTCAAAATTTTATTTTCTTTTGGTGGATATTT 4348  
Db 60279 TTTTCTCTCAATTTTAAATTTTGAAGTCAAAATTTTATTTTCTTTTGGTGGATATTT 60328

RESULT 5  
AAH13813  
ID AAH13813 standard; cDNA; 2044 BP.  
XX  
AC  
AC  
DT 26-JUN-2001 (first entry)  
XX  
DE Human cDNA sequence SEQ ID NO:10767.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX Homo sapiens.  
XX EP1074617-A2.  
XX  
XX 07-FEB-2001.  
XX 28-JUL-2000; 2000BP-00116126.  
XX 29-JUL-1999; 99JP-00248036.  
XX 27-AUG-1999; 99JP-00300253.  
XX 11-JAN-2000; 2000JP-00118776.  
XX 02-MAY-2000; 2000JP-00183767.  
XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.  
XX  
PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.  
PT  
PT  
PT  
PT  
PS Claim 8; SEQ ID NO 10767; 2537pp + Sequence Listing; English.  
XX  
CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
CC complementary strand of a polynucleotide which comprises one of the 5602  
CC nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
CC specification. The primer sets can be used in antisense therapy and in the  
CC gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893  
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
CC oligonucleotides, all of which are used in the exemplification of the  
CC present invention  
XX  
SQ Sequence 2044 BP; 723 A; 384 C; 442 G; 495 T; 0 U; 0 Other;  
Query Match 41.7%; Score 1836.4; DB 4; Length 2044;  
Best Local Similarity 94.3%; Pred. No. 0;  
Matches 1976; Conservative 0; Mismatches 1; Indels 119; Gaps 2;  
QY 13 CGCGGTTGCGCGTGGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72  
Db 67 CGCGGTTGCGCGTGGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 126  
QY 73 CGCGGCG 132  
Db 127 CGCGGCG 186  
QY 133 AAGGAAAAATCGTGCCTTCTTAATTTGAAAAATGATCAGTCGATCAGCCGAAATCATGCTGT 192  
Db 187 AAGGAAAAATCGTGCCTTCTTAATTTGAAAAATGATCAGTCGATCAGCCGAAATCATGCTGT 246  
QY 193 GTTAATCTGTAATCTTTCTGTAACCACTGATGATCAGTCAAGTCAAGTCAAGTCAAGTCAAG 252  
Db 247 GTTAATCTGTAATCTTTCTGTAACCACTGATGATCAGTCAAGTCAAGTCAAGTCAAGTCAAG 306  
QY 253 ATTTAAAGATTAATCTTAAGTATGTTACCTTTGTTAATGAGGAAAAATGAGGAAATGGCTT 312  
Db 307 ATTTAAAGATTAATCTTAAGTATGTTACCTTTGTTAATGAGGAAAAATGAGGAAATGGCTT 366  
QY 313 TTCCCGAACTTTGAAAGTGGGGGATGTTATTAATTTTGGAGTGTGTTGAGTAAATTCAG 372  
Db 367 TTCCCGAACTTTGAAAGTGGGGGATGTTATTAATTTTGGAGTGTGTTGAGTAAATTCAG 426  
QY 373 ATTAGATATGAGCCCTTTGTTGATGCTCTTCTTTGTTTATGATGCTCTCTGGAAGAACTGC 432  
Db 427 AATAGATATGAGCCCTTTGTTGATGCTCTTCTTTGTTTATGATGCTCTCTGGAAGAACTGC 486  
QY 433 TTTAAATCAAGCTATATATTGCACTTTGGAGGATTTACTGTAAACAATTTGACAGAGAAATG 492

Db 487 TTTAAATCAAGCTATATTGCAACTTGGAGGATTTACTGTAAACAATTTGGACAGAGAATG 546  
Qy 493 CACTCACCTTGTGTCATGTCATGTCAGTGAAGTTTACCATTAAAAACAATATGTGCACCTCATTTG 552  
Db 547 CACTCACCTTGTGTCATGTCATGTCAGTGAAGTTTACCATTAAAAACAATATGTGCACCTCATTTG 606  
Qy 553 TGGAGCTCCAAATGTTAAAGCAGAAATATTTTACTGAATTCCTGAAAGCAGTTCAGTCCAA 612  
Db 607 TGGAGCTCCAAATGTTAAAGCAGAAATATTTTACTGAATTCCTGAAAGCAGTTCAGTCCAA 666  
Qy 613 GAAGCAGCTCCCAAAATTCGAAGTTTATACCCACCTCTTGATGAACCATCTATTGGAAG 672  
Db 667 GAAGCAGCTCCCAAAATTCGAAGTTTATACCCACCTCTTGATGAACCATCTATTGGAAG 726  
Qy 673 TAAAAATGTTGATCTGTGTCAGGAGCGGAGGAAAGAAACAATCTTCAAGGGGAAACAATT 732  
Db 727 TAAAAATGTTGATCTGTGTCAGGAGCGGAGGAAAGAAACAATCTTCAAGGGGAAACAATT 786  
Qy 733 TATATTTTGAATGCCAAACAGCATAAAGAAATGAGTTCGCGAGTTGTCTTTTGGAGGTGG 792  
Db 787 TATATTTTGAATGCCAAACAGCATAAAGAAATGAGTTCGCGAGTTGTCTTTTGGAGGTGG 846  
Qy 793 GGAAGCTAGTTCATACAGAGAGATGAAGAAATGAGTTCGCGAGTTGTCTTTTGGAGGTGG 852  
Db 847 GGAAGCTAGTTCATACAGAGAGATGAAGAAATGAGTTCGCGAGTTGTCTTTTGGAGGTGG 906  
Qy 853 AAGCTGTGTTGATACAGGAATACAAACTCACAGACCTTAATTCCTGACTGTGAGAA 912  
Db 907 AAGCTGTGTTGATACAGGAATACAAACTCACAGACCTTAATTCCTGACTGTGAGAA 966  
Qy 913 GAAATGGATTCAGTCAATATGATATGCTCCAAAGCGCAAGGCTTAGACCTATTCCTGA 972  
Db 967 GAAATGGATTCAGTCAATATGATATGCTCCAAAGCGCAAGGCTTAGACCTATTCCTGA 1026  
Qy 973 AGCAGAAATGGATTTGGCGTGAATTTTCATGACTACAAAGAAATGATGTCCTCAGGG 1032  
Db 1027 AGCAGAAATGGATTTGGCGTGAATTTTCATGACTACAAAGAAATGATGTCCTCAGGG 1086  
Qy 1033 CCATCCAGTACAGGATTTAAAGCAACAACCTCAGGACCAAGCTTTTCAAGAGCGGTGC 1092  
Db 1087 CCATCCAGTACAGGATTTAAAGCAACAACCTCAGGACCAAGCTTTTCAAGAGCGGTGC 1146  
Qy 1093 AGTTGATGAAAACTAATGTCAGAGCGCCCGAGTGAACACTACACATACGTAGCTGACAC 1152  
Db 1147 AGTTGATGAAAACTAATGTCAGAGCGCCCGAGTGAACACTACACATACGTAGCTGACAC 1206  
Qy 1153 AGAATCAGACAGCAGATACATGGGATTTGAGTGAAGGCCAAAGAAATCAAGTCTC 1212  
Db 1207 AGAATCAGACAGCAGATACATGGGATTTGAGTGAAGGCCAAAGAAATCAAGTCTC 1266  
Qy 1213 CAAAAATGGAAACAAAAATTCAGAAATGCTTTTCAAGAGCGCACCCACTGTAAAGGAGCTTG 1272  
Db 1267 CAAAAATGGAAACAAAA----- 1282  
Qy 1273 CAAAAACAAGCTCTAATAATAATAGTATGATGATCAAAATACCTTTGGCTAAGATGAGAAATCCC 1332  
Db 1283 ----- 1282  
Qy 1333 AAACATACAGCTTTTACCACAACTAAATGCAAGTATATAATAAAGTAAAGATAGGCGCTTC 1392  
Db 1283 -----CACCACAACTAAATGCAAGTATATAATAAAGTAAAGATAGGCGCTTC 1328  
Qy 1393 TCAGCAGCAGCAGCAGCAACTCCATCAGAACTACTTTTACGCGCTCTTACC-AAAAAAGGG 1451  
Db 1329 TCAGCAGCAGCAGCAGCAACTCCATCAGAACTACTTTTACGCGCTCTTACC----- 1388  
Qy 1452 AAAGGGATGAAGAAAAATCAAGAAATGCTTTTCATGCAAAATCAGCAGAAATAGAACCTCTTT 1511  
Db 1389 AAAGGGATGAAGAAAAATCAAGAAATGCTTTTCATGCAAAATCAGCAGAAATAGAACCTCTTT 1448  
Qy 1512 GTTCTCTTTTAGAACAAAAACAACCTGCTACACCCCTCATTTGTGGAAAAAATAAGGAGCAGC 1571  
|||||

Db 1449 GTTCTCTTTTAGAACAAAAACAACCTGCTGTACACCCCTCATTTGTGGAAAAAATAAGGAGCAGC 1508  
Qy 1572 ATCTATCTCAGAAATGAGCCCTGTGGACACAAACTCAGACAATAACTTTATTATTACAGATACAG 1631  
Db 1509 ATCTATCTCAGAAATGAGCCCTGTGGACACAAACTCAGACAATAACTTTATTATTACAGATACAG 1568  
Qy 1632 ATTTAAAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGCTGCAGAAAAAGCTAAGAT 1591  
Db 1569 ATTTAAAAATCTATTGTGAAAAATTTCTGCCAGTAAATCTCATGCTGCAGAAAAAGCTAAGAT 1628  
Qy 1692 CAAATAAAAAAGGGAATGGAATGATGTGCCCATCAAGATGAAGTATTGGACACGTTAT 1751  
Db 1629 CAAATAAAAAAGGGAATGGAATGATGTGCCCATCAAGATGAAGTATTGGACACGTTAT 1688  
Qy 1752 TCAAGCAGACAAAAACAGAGTTTAGAAATTTGATGTGAAAGTTTCAAAAAACAGGAGGAAGATG 1811  
Db 1689 TCAAGCAGACAAAAACAGAGTTTAGAAATTTGATGTGAAAGTTTCAAAAAACAGGAGGAAGATG 1748  
Qy 1812 TCATGTTTGAAGAAAAAGGCCAAGGATGGATATAGAAAACAAATGACACTTTTCAGTGTGAAG 1871  
Db 1749 TCAATGTTTGAAGAAAAAGGCCAAGGATGGATATAGAAAACAAATGACACTTTTCAGTGTGAAG 1808  
Qy 1872 CAGTACCAGAAAAAGTAGCAAAATATCTCAAGAAAAATGAAATTTGGAAAGAACGTTGAACTCA 1931  
Db 1809 CAGTACCAGAAAAAGTAGCAAAATATCTCAAGAAAAATGAAATTTGGAAAGAACGTTGAACTCA 1868  
Qy 1932 AGGAAGACTCACCTATGCTCAGCTTAAAGAAATATCTAACTAATGACAACTTCAGGATGATA 1991  
Db 1869 AGGAAGACTCACCTATGCTCAGCTTAAAGAAATATCTAACTAATGACAACTTCAGGATGATA 1928  
Qy 1992 GTGAGATGCTTCCAAAAAGCTGTTATTGACTGAATTTAGATCAGTCTGGTGAATTAAGAACT 2051  
Db 1929 GTGAGATGCTTCCAAAAAGCTGTTATTGACTGAATTTAGATCAGTCTGGTGAATTAAGAACT 1988  
Qy 2052 CTACTTCCAGAAATCGTCTGCATTAATGATGATTTATGCTCAACTTAAATAATTTTC 2107  
Db 1989 CTACTTCCAGAAATCGTCTGCATTAATGATGATTTATGCTCAACTTAAATAATTTTC 2044  
|||||

## RESULT 6

AAH03660

ID AAH03660 standard; cDNA; 752 BP.

XX AAH03660;

XX AC

XX XX

XX 26-JUN-2001 (first entry)

XX DE Human cDNA clone (5'-primer) SEQ ID NO:495.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX XX

XX 07-FEB-2001.

XX XX

XX 28-JUL-2000; 2000EP-00116126.

XX XX

XX 29-JUL-1999; 99JP-00248036.

XX PR 27-AUG-1999; 99JP-00300253.

XX PR 11-JAN-2000; 2000JP-00118776.

XX PR 02-MAY-2000; 2000JP-00183767.

XX PR 09-JUN-2000; 2000JP-00241899.

XX XX

XX (HELI-) HELIX RES INST.

XX XX

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX XX

XX WPI; 2001-318749/34.

XX XX

XX PI Primer sets for synthesizing polynucleotides, particularly the 5602 full-

PT length cDNAs defined in the specification, and for the detection and/or  
PT diagnosis of the abnormality of the proteins encoded by the full-length  
PT cDNAs.

PS Claim 1; SEQ ID NO 495; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesising 5602 full-  
CC length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dr primer and an oligonucleotide complementary to the  
CC complementary strand of a polynucleotide which comprises one of the 5602  
CC nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
CC specification. The primer sets can be used in antisense therapy and in  
CC gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893  
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
CC oligonucleotides, all of which are used in the exemplification of the  
CC present invention

XX SQ Sequence 752 BP; 212 A; 158 C; 177 G; 202 T; 0 U; 3 Other;

Query Match 15.2%; Score 671; DB 4; Length 752;

Best Local Similarity 99.4%; Pred. No. 1.1e-131;

Matches 682; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 13 CGCGGTTGACGTCGCGCCCGCCGAGCGGCGGAGCGGAGTGGAAACTGTCGCCCG 72

DB 67 CGCGGTTGACGTCGCGCCCGCCGAGCGGCGGAGCGGAGTGGAAACTGTCGCCCG 126

QY 73 CGCGGCGCGGAGGAGAGAACATACAGACTTTTGTGACTGGCGTTGAGTACGTTGG 132

DB 127 CGCGGCGCGGAGGAGAGAACATACAGACTTTTGTGACTGGCGTTGAGTACGTTGG 186

QY 133 AAGGAAACTGTCGCACTTCTAATGAAATGATCATGTCAGTCAGCGGAAATCATGCTCT 192

DB 187 AAGGAAACTGTCGCACTTCTAATGAAATGATCATGTCAGTCAGCGGAAATCATGCTCT 246

QY 193 GTTAACTGCTAACTTTCTGTAAACCACTGAGTCAAAAGATGAAATCCCTGATTGAC 252

DB 247 GTTAACTGCTAACTTTCTGTAAACCACTGAGTCAAAAGATGAAATCCCTGATTGAC 306

QY 253 ATTTAAAGATAATCTTAAGTATGGTACCTTTGTTAATGAGGAAAAAATGAGAATGGCTT 312

DB 307 ATTTAAAGATAATCTTAAGTATGGTACCTTTGTTAATGAGGAAAAAATGAGAATGGCTT 366

QY 313 TTCCCGAATTTGAACTCGGGGATGGTATTACTTTTGGAGTCTTGGAGTAAATTCAG 372

DB 367 TTCCCGAATTTGAACTCGGGGATGGTATTACTTTTGGAGTCTTGGAGTAAATTCAG 426

QY 373 AATAGAGTATGAGCTTTGGTTCATGCTCTTCTGTTAGATGTCCTGGGAAAACTGC 432

DB 427 AATAGAGTATGAGCTTTGGTTCATGCTCTTCTGTTAGATGTCCTGGGAAAACTGC 486

QY 433 TTTAAATCAAGCTATATTTGCAACTTGGAGATTATCTGTAAACAATTTGGACAGAAATG 492

DB 487 TTTAAATCAAGCTATATTTGCAACTTGGAGATTATCTGTAAACAATTTGGACAGAAATG 546

QY 493 CACTCACCTTGTATGCTGTAAGTACCTTAAACCAATATGTCATCTATTG 552

DB 547 CACTCACCTTGTATGCTGTAAGTACCTTAAACCAATATGTCATCTATTG 606

QY 553 TGGACGTCCTCAATTTGTAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAA 612

DB 607 TGGAGCTCCAAATTTGTAAGCCAGAAATATTTTACTGAATTCCTGAAAGCAGTTTCAGTCCAA 666

QY 613 GAAGCAGGCTCCACAAATTTGAAGTTTATCCCACTCTTGTATGAAACCATCTTATTTGG-AA 671

DB 667 GAAGCAGGCTTCACAAATTTGAAGTTTATCCCACTCTTGTATGAAACCATCTTATTTGGAAA 726

QY 672 GTAAAAATGTTGATCTGTCTGTCAGGACGG 697

DB 727 GTAAAAATGTTGATCTGTCTGTCAGGACGG 752

RESULT 7

ABZ71727

ID ABZ71727 standard; cDNA; 622 BP.

AC ABZ71727;

XX 01-APR-2003 (first entry)

DE Human cancer-associated gene SEQ ID NO 34.

XX Human; cancer; stomach cancer; cytostatic; gene; sb.

XX Homo sapiens.

XX WO200283899-A1.

XX 24-OCT-2002.

XX 28-MAR-2002; 2002WO-JP003038.

XX 10-APR-2001; 2001JP-00112039.

XX 21-SEP-2001; 2001JP-00290193.

XX (TAKA-) TAKARA BIO INC.

XX Yoshikawa Y, Okamoto S, Oura T, Mineno J, Asada K, Kato I;

XX Inoue H, Mori M;

XX WPI; 2003-093022/08.

XX Measuring changes in expression of 264 cancer associated genes for

XX detection of stomach cancer and screening of potential anticancer agents.

XX Claim 1; Page 114; 266pp; Japanese.

XX The invention relates to a method for the detection of cancer in which a

XX change in the expression of 1 or more of 264 specified cancer associated

XX genes, ABZ71694-ABZ71957, or of sequences at least 80% homologous to them

XX in the specimen tissue as compared to normal tissue is observed. The

XX genes are used in detection, diagnosis and treatment of cancer,

XX especially of stomach cancer. The present sequence is that of a cancer

XX associated polynucleotide of the invention

XX SQ Sequence 622 BP; 230 A; 131 C; 121 G; 138 T; 0 U; 2 Other;

Query Match 12.8%; Score 564.6; DB 8; Length 622;

Best Local Similarity 97.6%; Pred. No. 3.2e-109;

Matches 604; Conservative 0; Mismatches 11; Indels 4; Gaps 3;

QY 840 TTTTGGCTCCGGGAACGCTGTGTGTGATACAGGAATAACAACTCACAGACTTTAATTC 899

DB 1 TTTTGGCTCCGGGAACGCTGTGTGTGATACAGGAATAACAACTCACAGACTTTAATTC 60

QY 900 CTGACTGTGAGAGAAATGGATTCTAGTCAATAATGGATATGCTCAAAGGCAAGTCTTAA 959

DB 61 CTGACTGTGAGAGAAATGGATTCTAGTCAATAATGGATATGCTCAAAGGCAAGTCTTAA 120

QY 960 GACCTATTCTTCAAGCAGAAATTTGATTTGGCGGTGATTTTCATGACTACAAAGATTACT 1019

DB 121 GACCTATTCTTCAAGCAGAAATTTGATTTGGCGGTGATTTTCATGACTACAAAGATTACT 180

QY 1020 GTGATCCTCAGGGCCATCCAGTACAGGATTAAGACAACAACTCCAGGACCAAGCCTTT 1079



RESULT 9  
 AAI28794/c  
 ID AAI28794 standard; cDNA; 646 BP.  
 XX  
 AC  
 XX AAI28794;  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE Colon tumour related determined cDNA sequence for clone 25908.  
 XX  
 KW Human; immunotherapy; diagnosis; colon cancer; colon tumour; immunogenic;  
 KW gene therapy; vaccine; colonic cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200149716-A2.  
 XX  
 PD 12-JUL-2001.  
 XX  
 PF 29-DEC-2000; 2000WO-US035596.  
 XX  
 PR 30-DEC-1999; 99US-00476296.  
 PR 10-JAN-2000; 2000US-00480321.  
 PR 15-FEB-2000; 2000US-00504629.  
 PR 06-MAR-2000; 2000US-00519444.  
 PR 19-MAY-2000; 2000US-00575251.  
 PR 29-JUN-2000; 2000US-00609448.  
 PR 28-AUG-2000; 2000US-00649811.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;  
 PI King GE, Wang T, Jiang Y;  
 XX  
 DR WPI; 2001-441847/47.  
 XX  
 PT Colon tumor associated proteins and nucleic acids useful for the  
 PT prevention, diagnosis and treatment of colonic cancer.  
 XX  
 PS Claim 25; Page 228; 472pp; English.  
 XX  
 CC The present invention describes colon tumour associated proteins (I) and  
 CC the polynucleotides (II) that encode them. (I) have cytostatic activity.  
 CC (I) and (II) can be used in gene therapy and vaccine production. (I) and  
 CC (II) may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate colon tumour associated protein (TCAP)  
 CC expression, such as colonic cancer. For example, (I) and (II) may be used  
 CC to treat disorders associated with decreased expression by rectifying  
 CC mutations or deletions in a patient's genome that affect the activity of  
 CC TCAPs by expressing inactive proteins or to supplement the patients own  
 CC production of them. Additionally, (II) may be used to produce the TCAP  
 CC proteins, by inserting the nucleic acids into a host cell culturing the  
 CC cell to express the protein. (II) and its complementary sequences may  
 CC also be used as DNA probes in diagnostic polymerase chain reaction (PCR)  
 CC and hybridisation assays to detect and quantitate the presence of similar  
 CC nucleic acids in samples, and therefore which patients may be in need of  
 CC restorative therapy. (I) may also be used as antigens in the production  
 CC of antibodies against TCAPs and in assays to identify modulators of TCAP  
 CC expression and activity. Anti-(I) antibodies and antagonists may also be  
 CC used to down regulate TCAP expression and activity. The anti-(I)  
 CC antibodies may also be used as diagnostic agents for detecting the  
 CC presence of TCAPs in samples (e.g. by enzyme linked immunosorbant assay  
 CC (ELISA)). AAI28460 to AAI29512 and AAM24494 to AAM24523 represent  
 CC nucleotide and amino acid sequences given in the exemplification of the  
 CC present invention  
 XX  
 SQ Sequence 646 BP; 239 A; 124 C; 95 G; 188 T; 0 U; 0 Other;  
 Query Match 12.5%; Score 552.2; DB 4; Length 646;  
 Best Local Similarity 93.0%; Pred. No. 1.4e-106;  
 Matches 609; Conservative 4; Mismatches 32; Indels 10; Gaps 3;  
 QY 3419 AAAAGCTTCTCAGCCTTCCTAGGGAACAGAAATGGGTGAGCCAATCTGCAATTTCTACT 3478

Db ||||| 645 AAAGCTTCTCAGCCTTCCTAGGGAACAGAAATGGGAGCCCAATCTGCAATTTCTACT 586  
 QY ||||| 3479 ACAGGCATTGAGACCCAGTTAGATTATTTGAAATATTTATAGAGAGTTATGAACTTAATTT 3538  
 Db ||||| 585 ACAGGCATTGAGACCCAGTTAGATTATTTGAAATATTTATAGAGAGTTATGAACTTAATTT 526  
 QY ||||| 3539 ATGATAGTGGTATGACATTGGATAGAACATGGGATACCTTTAGAACTAGAAATTTGACAGGC 3598  
 Db ||||| 525 ATGATAGTGGTATGACATTGGATAGAACATGGGATACCTTTAGAACTAGAAATTTGACAGGC 466  
 QY ||||| 3599 ATATTAGTTGATGAAATGGAGTCATTTCAGTCTCTTTAATAGCCATGTATCATTAATTTACCA 3658  
 Db ||||| 465 ATATTAGTTGATGAAATGGAGTCATTTCAGTCTCTTTAATAGCCATGTATCATTAATTTACCA 406  
 QY ||||| 3659 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATAATGGACAGATT 3718  
 Db ||||| 405 AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATAATGGACAGATT 346  
 QY ||||| 3719 AATATTGTTTCTGTCATGCCACAAATCCCTTTTCTAAGGAAGACTGCCCTACTATAGCAG 3778  
 Db ||||| 345 AATATTGTTTCTGTCATGCCACAAATCCCTTTTCTAAGGAAGACTGCCCTACTATAGCAG 286  
 QY ||||| 3779 TTTTATATTTTCTCAATTTATGAATATATAATGAATGAGGAGTTCTGGTACCTCTGTCCTTT 3838  
 Db ||||| 285 TTTTATATTTTCTCAATTTATGAATATATAATGAATGA- GAGTTCTGTTGACCTCTGTCCTTT 227  
 QY ||||| 3839 ACAATATTTGGTGTGTCAGATATTTTCCCTTTTAAACMTTCCCAATTCGGGTGTGT 3898  
 Db ||||| 226 ACAATATTTGG- - -TGTGTGTCAGTATTTTTCCTTTTAAACCAATTCCTGTTGTGTAG 171  
 QY ||||| 3899 AGTGGATGTTTCCATTTTGGGTTTAAATTTGTATATCCCTGATAGCTATAATTTGGTCTAT 3958  
 Db ||||| 170 TGATG- - -TTTCAATTTTGGTTTAAATTTGTATATCCCTGATAGCTATAATTTGGTCTAT 116  
 QY ||||| 3959 AGAAATTTCTTTATACATTTAGATGCAAGTCTCTTTGYCGGATATACGTATTCAGATATTTA 4018  
 Db ||||| 115 AGAAATTTCTTTATACATTTAGATGCAATTTCTTTGTCGGATATATGTATTGAGATATTTA 56  
 QY ||||| 4019 CACCTAGTCTGTGGCTTGACGTCTTTTCTTTATGCTCTTTTGTAGTAAAGATTTT 4073  
 Db ||||| 55 CACCTAGTCTGTGGCTTGACGTCTTTTCTTTATGCTCTTTTGTAGTAAAGATTTT 1  
 RESULT 10  
 ABZ32980/c  
 ID ABZ32980 standard; cDNA; 646 BP.  
 XX  
 AC ABZ32980;  
 XX  
 DT 30-JAN-2003 (first entry)  
 XX  
 DE Human colon tumour cDNA clone 25908 SEQ ID NO:343.  
 XX  
 KW Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;  
 KW tumour; immune response; immunostimulant; cytostatic; vaccine; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283070-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 09-APR-2002; 2002WO-US011475.  
 XX  
 PR 10-APR-2001; 2001US-00833263.  
 PR 03-AUG-2001; 2001US-00922217.  
 PR 19-DEC-2001; 2001US-00025380.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Xu J, Lodes MJ, Secretist H, Benson DR, Meagher MJ, Stolk JA;  
 PI Wang T, Jiang Y, Smith CL, King GE, Wang A, Clapper JD, Skeiky YAW;

PI Fanger GR, Vedwick TS, Carter D;  
XX WPI; 2003-067548/06.  
XX New polynucleotide, useful for the preparation of a composition for  
PT stimulating an immune response against, or treating, cancer.  
XX  
XX Example 1; Page 231; 537pp; English.  
XX  
XX The present invention describes compounds (I) for the immunotherapy and  
CC diagnosis of colon cancer. Also described: (1) a method for detecting the  
CC presence of cancer in a patient; (2) a method for stimulating and/or  
CC expanding T cells specific for a tumour protein; (3) an isolated T cell  
CC population comprising T cells prepared by the method of (2); (4) a method  
CC for stimulating an immune response in a patient; (5) a method for  
CC treating cancer in a patient; and (6) a method for inhibiting the  
CC development of cancer in a patient. (I) have immunostimulant and  
CC cytostatic activities and can be used in vaccines. AB232646 to AB233725  
CC and ABP55343 to ABP55391 represent human colon cancer/tumour related  
CC sequences used in the exemplification of the present invention  
XX  
SQ Sequence 646 BP; 239 A; 124 C; 95 G; 188 T; 0 U; 0 Other;  
  
Query Match 12.5%; Score 552.2; DB 8; Length 646;  
Best Local Similarity 93.0%; Pred. No. 1.4e-106;  
Matches 609; Conservative 4; Mismatches 32; Indels 10; Gaps 3;  
  
QY 3419 AAAAGCTTCTCAGCCTTCTCCTAGGACAGAAATTTGGTGAGCCAAATCTGCAATTTCTACT 3478  
DB 645 AAAGCTTCTCAGCCTTCTCCTAGGACAGAAATTTGGTGAGCCAAATCTGCAATTTCTACT 586  
  
QY 3479 ACAGGCATTGAGACCACTAGATTATTTGAATATTTAGAGAGTTATGACACTTAAAT 3538  
DB 585 ACAGGCATTGAGACCACTAGATTATTTGAATATTTAGAGAGTTATGACACTTAAAT 526  
  
QY 3539 ATGATAGTGTATGACATTCGATAGAACATCGGATCTTTAGAGTAGAATTCAGAGGGC 3598  
DB 525 ATGATAGTGTATGACATTCGATAGAACATCGGATCTTTAGAGTAGAATTCAGAGGGC 466  
  
QY 3599 ATATTAGTTCATGAATGGAGTCAATTTGAGTCTTCTTAATAGCCATGATCATTAATACCA 3658  
DB 465 ATATTAGTTCATGAATGGAGTCAATTTGAGTCTTCTTAATAGCCATGATCATTAATACCA 406  
  
QY 3659 AGTGAAGCTGGTGGAAACATATGCTCTCCATTTTACAGTTAAGGAATATATGACAGATT 3718  
DB 405 AGTGAAGCTGGTGGAAACATATGCTCTCCATTTTACAGTTAAGGAATATATGACAGATT 346  
  
QY 3719 AATATTGTTTCTGTCATGCCCAATCCCTTTCTAAGGAAGACTGCGCTTACTATAGCAG 3778  
DB 345 AATATTGTTTCTGTCATGCCCAATCCCTTTCTAAGGAAGACTGCGCTTACTATAGCAG 286  
  
QY 3779 TTTTATATTGTCATTTATGATTAATGAATGAATGAGGAGTCTGTTACCTCCTGCTTTT 3838  
DB 285 TTTTATATTGTCATTTATGATTAATGAATGAATGAGGAGTCTGTTACCTCCTGCTTTT 227  
  
QY 3839 ACAATATTGGGTGTTGTCAGTATTTTCCCTTTTAAACMTTCCCAATTCGGGTGTT 3898  
DB 226 ACAATATTGG-----TGTTGTCAGTATTTTCCCTTTTAAACATTCCAATCGGTGTAG 171  
  
QY 3899 AGTGGATGTTCCATTTGGGTTTTAATTTGATATATCCCTGATAGCTATTAATTTGGGTAT 3958  
DB 170 TGATG-----TTTCAITTTGGTTTTAATTTGATATATCCCTGATAGCTATTAATTTGGGTAT 116  
  
QY 3959 AGAATTTCTTATACATTTAGATGCAAGTCTCTTGGATATAGCTATGATGATTA 4018  
DB 115 AGAATTTCTTATACATTTAGATGCAAGTCTCTTGGATATAGCTATGATGATTA 56  
  
QY 4019 CACCTAGTCTGCGCTGACTGTTTCTTTATGCTCTTTTATGCTTTTATGCTTTTATGCTTTT 4073  
DB 55 CACCTAGTCTGCGCTGACTGTTTCTTTATGCTCTTTTATGCTTTTATGCTTTTATGCTTTT 1

ABQ57768/c  
ID ABQ57768 standard; cDNA; 561 BP.  
XX  
AC ABQ57768;  
XX  
DT 02-AUG-2002 (first entry)  
XX  
DE Human colon cancer related nucleotide sequence SEQ ID NO:1463.  
XX  
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
KW genetic analysis; diagnostic; antisense therapy; gene; ss.  
OS Homo sapiens.  
XX  
PN WO200229086-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 02-OCT-2001; 2001WO-US030732.  
XX  
PR 02-OCT-2000; 2000US-0237271P.  
XX  
PA (FARB ) BAYER CORP.  
XX  
PI Burgess C, Astle JH, Carroll B, Catino TJ, Dwivedi P, Molino GA;  
PI Thiagalingam A, Lewis ME;  
XX  
PS WPI; 2002-426115/45.  
XX  
PT New isolated nucleic acid that is differentially expressed in cancer  
PT tissues useful for determining the presence of colon cancer in a cell or  
XX tissue type, and in antisense therapy.  
XX  
PS Claim 1; Fig 1; 796pp; English.  
XX  
CC ABQ5306 to ABQ60787 represent isolated nucleic acids (I) differentially  
CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins  
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be  
CC used in antisense therapy. An antibody immunoreactive with a polypeptide  
CC encoded by (I) is useful for detecting cancer in a patient sample, and  
CC for detecting the presence or absence of a polynucleotide encoded by a  
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived  
CC from (I) can be used for determining the presence of a nucleic acid which  
CC hybridises to (I), and for determining the phenotype of cells in a sample  
CC of cells from a patient. (I) is useful for determining the presence of  
CC colon cancer in a cell or tissue type, for determining the presence or  
CC state of other type of cancer, in antisense therapy, to generate  
CC macroarrays on a solid surface, to identify a chromosome on which the  
CC corresponding gene resides, and in tissue profiling, forensics, genetic  
CC analysis, mapping and diagnostic applications. (I) can be used to raise  
CC antibodies, and to screen for peptide analogues and antagonists  
XX  
SQ Sequence 561 BP; 151 A; 120 C; 102 G; 187 T; 0 U; 1 Other;  
  
Query Match 12.4%; Score 544.8; DB 6; Length 561;  
Best Local Similarity 99.3%; Pred. No. 4.8e-105;  
Matches 557; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
  
QY 483 CAGAAGAAATGCACCTCA-CCTTGTCTCATGGTATCAGTGAAAGTTACCAATTAACAATATGT 541  
DB 561 CAGAAGAAATGCACCTCCTCCCTTGTCTCATGGTATCAGTGAAAGTTACCAATTAACAATATGT 502  
  
QY 542 GCACCTCATTTGTGGAGCTCCAAATTTGAAAGCCAGAAATATTTTACTCAATTTCTTGAAGCA 601  
DB 501 GCACCTCATTTGTGGAGCTCCAAATTTGAAAGCCAGAAATATTTTACTGAAATTTCTTGAAGCA 442  
  
QY 602 GTTCACTCCAAGAAGCAGCCTCCACAAATTTGAAAGTTTACCACCTCTTGTATGAACCA 661  
DB 441 GTTCACTCCAAGAAGCAGCCTCCACAAATTTGAAAGTTTACCACCTCTTGTATGAACCA 382  
  
QY 662 TCTATTGGAAGTAAAAATGTTGATCTGTGAGACGCGCAGGAAAGAAACAAATCTTCAAA 721  
DB 381 TCTATTGGAAGTAAAAATGTTGATCTGTGAGACGCGCAGGAAAGAAACAAATCTTCAAA 322

Qy	722	GGGAAACATTTATATTTTGAATGCGAAACAGCATAGAATTTGAGTTCGCGAGTTGTC	781
Db	321	GGGAAACATTTATATTTTGAATGCGAAACAGCATAGAATTTGAGTTCGCGAGTTGTC	262
Qy	782	TTTGGAGTGGGGAACGTAGGTTGATACAGAGAGAAATGAAGAAGCAATAAATTTCTTT	841
Db	261	TTTGGAGTGGGGAACGTAGGTTGATACAGAGAGAAATGAAGAAGCAATAAATTTCTTT	202
Qy	842	TTGGCTCCGGGAACGTGTTGATACAGGAAATAACAACTCACAGACCTTAATTCCT	901
Db	201	TTGGCTCCGGGAACGTGTTGATACAGGAAATAACAACTCACAGACCTTAATTCCT	142
Qy	902	GACTGTCAGAGAAATGGATTCAGTCAATATATGATATGCTCCAAAGCAAGTCTTAGA	961
Db	141	GACTGTCAGAGAAATGGATTCAGTCAATATATGATATGCTCCAAAGCAAGTCTTAGA	82
Qy	962	CCTATTCTCGAAGCAGAGAAATGGATTGCGGTGATTTTCATGACTACAAAGAATTACTGT	1021
Db	81	CCTATTCTCGAAGCAGAGAAATGGATTGCGGTGATTTTCATGACTACAAAGAATTACTGT	22
Qy	1022	GATCCTCAGGGCCATCCCACT 1042	
Db	21	GATCCTCAGGGCCATCCCACT 1	
RESULT 12			
ABL38148			
XX	ID	ABL38148 standard; cDNA; 544 BP.	
AC	AC	ABL38148;	
XX	XX		
DT	DT	08-APR-2002 (first entry)	
DE	DE	Human colon tumour antigen polynucleotide SEQ ID NO:1737.	
XX	XX		
KW	KW	Human; colon cancer; colon tumour antigen; cytostatic; vaccine;	
KW	KW	colon tumour metastatic antigen; diagnosis; gene; ss.	
XX	XX		
OS	OS	Homo sapiens.	
PN	PN	WO2001196388-A2.	
XX	XX		
PD	PD	20-DEC-2001.	
XX	XX		
PF	PF	08-JUN-2001; 2001WO-US018557.	
XX	XX		
PR	PR	09-JUN-2000; 2000US-0210899P.	
XX	XX	20-FEB-2001; 2001US-0270216P.	
XX	XX		
PA	PA	(CORI-) CORIXA CORP.	
XX	XX		
PI	PI	Jiang Y, Harlocker SL, Secretist H;	
XX	XX		
DR	DR	WPI; 2002-114514/15.	
XX	XX		
PT	PT	Novel isolated colon tumor polynucleotide differentially expressed in	
PT	PT	colon tumor or colon metastatic tumor and polypeptides encoded by them,	
PT	PT	useful for inhibiting development of cancer in patient.	
XX	XX		
PS	PS	Claim 1; SEQ ID NO 1737; 105pp; English.	
XX	XX		
CC	CC	ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)	
CC	CC	which were isolated from human colon tumour and colon metastatic tumour	
CC	CC	cDNA libraries. (I) have cytostatic activity and can be used in vaccine	
CC	CC	production. (I) can be used for stimulating and/or expanding T cells	
CC	CC	specific for a tumour protein on contact with the T cells. They are also	
CC	CC	useful for inhibiting the development of cancer in a patient. (I) can be	
CC	CC	used as probes or primers for nucleic acid hybridisation, for preparing	
CC	CC	mutant species primers, or primers for use in genetic constructions. (I)	
CC	CC	can be used in the diagnosis of a colon tumour	
XX	XX		
XX	XX	Sequence 544 BP; 173 A; 94 C; 125 G; 150 T; 0 U; 2 Other;	

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-  
PT length cDNAs defined in the specification, and for the detection and/or  
PT diagnosis of the abnormality of the proteins encoded by the full-length  
PT cDNAs.  
XX  
XX  
PS Claim 3; SEQ ID NO 6023; 2537pp + Sequence Listing; English.  
XX  
CC The present invention describes primer sets for synthesizing 5602 full-  
CC length cDNAs defined in the specification. Where a primer set comprises:  
CC (a) an oligo-dT primer and an oligonucleotide complementary to the  
CC complementary strand of a polynucleotide which comprises one of the 5602  
CC nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in the  
CC specification. The primer sets can be used in antisense therapy and in  
CC gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893  
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent  
CC oligonucleotides, all of which are used in the exemplification of the  
CC present invention  
XX  
SQ Sequence 543 BP; 128 A; 111 C; 83 G; 219 T; 0 U; 2 Other;

Query Match 11.8%; Score 520.4; DB 4; Length 543;  
Best Local Similarity 98.3%; Pred. No. 6.8e-100;  
Matches 535; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 1564 GGAGCAGCATCTATCTGAGAATGAGCGCTGTGGACACAAACTCAGACAAATTAATTAC 1623  
DB |||||  
QY 1624 AGATCAGATTTAAATCTATCTGAGAAATCTGCCAGTAATCTCATGCTGCACAAA 1683  
DB |||||  
QY 1684 GCTAAGATCAATATAAAGGAAATGGATGATGTGGCCATAGAGATGAAGTATTGGA 1743  
DB |||||  
QY 1744 ACAGTTATTCAAGGACACAAAACACAGATTAGAAATTTGATGTAAGATTTCAAAAACAGGA 1803  
DB |||||  
QY 1804 GGAAGATGTCATGTTAGAAAAGGCCAAGGATGGATATAGAAACAATGACACTTTCAG 1863  
DB |||||  
QY 304 GGAAGATGTCATGTTAGAAAAGGCCAAGGATGGATATAGAAACAATGACACTTTCAG 245  
QY 1864 TGATGAAGCAGTACACAGAAGTAGCAAAATATCTCAAGAAAATGAAATTTGGGAAGAACG 1923  
DB |||||  
QY 1924 TGAACCTCAAGGAAGACTCACTATGTCAGCTAAGAAAATATCTAACAATGACAAACTTCA 1983  
DB |||||  
QY 184 TGAACCTCAAGGAAGACTCACTATGTCAGCTAAGAAAATATCTAACAATGCGCAACTTCA 125  
QY 1984 GGATGATAGTAGAGATGCTTCCAAAAAGCTGTTATTGACTGAATTTAGATCACTGGTAT 2043  
DB |||||  
QY 124 GGATGATAGTAGAGATGCTTCCAAAAAGCTGTTATTGACTGAATTTAGATCACTGGTAT 65

QY 2044 TAAAAACTCTACTTCCAGAAATCCGCTCTGCATAAATGATGATTATGGTCAACTAAAAA 2103  
DB |||||  
QY 2104 TTTC 2107  
DB |||||  
DB 4 TTTC 1  
RESULT 14  
ADST71935  
ID ADS71935 standard; cDNA; 587 BP.  
XX  
XX ADS71935;  
XX  
XX 18-NOV-2004 (first entry)  
XX  
XX Human kidney tumour specific cDNA, SEQ ID 532.  
XX  
XX Human; ss; kidney tumour; kidney cancer; cytostatic; gene therapy;  
KW T-cell; immune response.  
XX  
XX Homo sapiens.  
XX  
XX US2003109434-A1.  
PN  
PD 12-JUN-2003.  
XX  
XX 19-MAR-2002; 2002US-00102524.  
PF  
XX 19-MAR-2001; 2001US-0277245P.  
PR  
XX 21-DEC-2001; 2001US-0343340P.  
XX  
XX (CORI-) CORIXA CORP.  
PA  
XX  
XX  
PI Algate PA, Mannion J, Gaiger A, Gordon B, Harlocker SL;  
XX  
XX WPI; 2002-759855/82.  
XX  
XX New isolated polynucleotides and polypeptides, useful for detecting the  
PT presence of, and treating cancer, particularly kidney cancer by  
PT stimulating T-cells specific for a tumor protein, and stimulating immune  
PT response in a patient.  
XX  
XX Claim 1; SEQ ID NO 532; 78pp; English.  
PS  
XX  
XX The invention relates to a new isolated polynucleotide (a Human kidney  
CC tumour specific cDNA) comprising any one of the 1855 sequences identified  
CC in the specification (or their complements, degenerate variants,  
CC sequences consisting of at least 20 contiguous residues them, sequences  
CC that hybridise to them under highly stringent conditions or sequences  
CC having at least 75 or 90% sequence identity to the 1855 sequences. Also  
CC included are detecting/determining the presence of cancer in a patient,  
CC stimulating an immune response in a patient; treating kidney cancer in a  
CC patient, an isolated polypeptide encoded by one of the 1855 sequences, an  
CC expression vector comprising the polynucleotide operably linked to an  
CC expression control sequence, a host cell transformed/transfected with the  
CC vector, an isolated antibody (or its antigen-binding fragment) that  
CC specifically binds to the protein, a fusion protein comprising at least  
CC one of the proteins, stimulating and/or expanding T-cells specific for a  
CC tumour protein, an isolated T-cell population comprising the T-cells, a  
CC composition comprising a first component (such as a carrier or  
CC immunostimulant) and a second component (comprising one of the  
CC polynucleotides, the polypeptides, an antibody, T-cell or an antigen-  
CC presenting cell that expresses the polynucleotide) and a diagnostic kit  
CC comprising at least one of the oligonucleotides, or at least one antibody  
CC and a detection reagent comprising a reporter group. The polynucleotides,  
CC polypeptides, antibodies and antigen-presenting cells are useful for  
CC detecting the presence of, and treating cancer, particularly kidney  
CC cancer by stimulating and/or expanding T-cells specific for a tumour  
CC protein, and stimulating immune response in a patient. The present  
CC sequence is one of the Human kidney tumour specific cDNAs. Note: The

CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from USPTO  
CC at seqdata.uspto.gov/sequence.html?DocID=20030109434.  
XX SQ Sequence 587 BP; 177 A; 86 C; 114 G; 210 T; 0 U; 0 Other;  
Query Match 11.5%; Score 507; DB 7; Length 587;  
Best Local Similarity 93.6%; Pred. No. 4.7e-97;  
Matches 559; Conservative 4; Mismatches 24; Indels 10; Gaps 3;  
QY 3419 AAAAGCTTCTCAGCCTTCCTAGGAAACAGAAATGGGTGAGCCAAATCTGCAATTTCTACT 3478  
Db 1 AAAAGCTTCTCAGCCTTCCTAGGAAACAGAAATGGGTGAGCCAAATCTGCAATTTCTACT 60  
QY 3479 ACAGGCAATTCAGACCAAGTAGATATTGAAATATATAGAGAGTTATGAACACTTAAATTT 3538  
Db 61 ACAGGCAATTCAGACCAAGTTAGATATTGAAATATATATAGAGAGTTATGAACACTTAAATTT 120  
QY 3539 ATGATAGTGTATGACATTTGGATAGAACATGGGATACTTTAGAAGTAGAATTGACAGGGC 3598  
Db 121 ATGATAGTGTATGACATTTGGATAGAACATGGGATACTTTAGAAGTAGAATTGACAGGGC 180  
QY 3599 ATATTAGTTGATGAAATGAGTCAATTTGAGTCTCTTTTAATAGCCATGTATCATAATTACCA 3658  
Db 181 ATATTAGTTGATGAAATGAGTCAATTTGAGTCTCTTTTAATAGCCATGTATCATAATTACCA 240  
QY 3659 AGTGAAGCTGGTGGAAACATATGCTCTCCATTTTACAGTTAAGGAATATATGACACAGATT 3718  
Db 241 AGTGAAGCTGGTGGAAACATATGCTCTCCATTTTACAGTTAAGGAATATATGACACAGATT 300  
QY 3719 AATATTGTTCTGTCATGCCCAATCCCTTCTAAGGAAGACTGCCCTACTATAGCAG 3778  
Db 301 AATATTGTTCTGTCATGCCCAATCCCTTCTAAGGAAGACTGCCCTACTATAGCAG 360  
QY 3779 TTTTATATTGTCATTTATGAATATGAATGAGGAGTTCCTGTACTCTCTGTCCTTT 3838  
Db 361 TTTTATATTGTCATTTATGAATATGAATGA-TGA-GAGTCTGGTACTCTGTCCTTT 419  
QY 3839 ACAATATTGGGTGTTGTCAGTATTTTCCCTTTTAAACMTTCCCAATTCGGGTGTT 3898  
Db 420 ACAATATTGG-----TGTGTGTCAGTATTTTCCCTTTTAAACCATTCCAATTCGGGTGTTAG 475  
QY 3899 AGGTGATGTTTCCATTTGGGTTTTAAATTTGTATATCCCTGTAGTCTATAATTTGGGTGAT 3958  
Db 476 TGAAG-----TTTCATTTTGGTTTTAAATTTGTATATCCCTGTAGTCTATAATTTGGGTGAT 530  
QY 3959 AGAAATTCCTTATACATCTAGATGCAAGTCTCTTGGYCGGATATACGTATTGAGATA 4015  
Db 531 AGAAATTCCTTATACATCTAGATGCAAGTCTCTTGGYCGGATATATGATTGAGATA 587  
RESULT 15  
ID ACH26622  
XX ACH26622 standard; cDNA; 468 BP.  
AC ACH26622;  
XX  
XX 13-OCT-2003 (first entry)  
DT Human adult ovary cDNA #5002.  
DE Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
KW Genome mapping; biodiversity; genetic disorder.  
XX Homo sapiens.  
XX US2003073623-A1.  
PN 17-APR-2003.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
XX 30-JUL-2001; 2001US-00918995.  
PR

XX (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
PI WPI; 2003-615964/58.  
XX  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 13834; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX  
XX Sequence 468 BP; 136 A; 86 C; 102 G; 141 T; 0 U; 3 Other;  
SQ  
Query Match 9.2%; Score 405.6; DB 9; Length 468;  
Best Local Similarity 97.4%; Pred. No. 1.2e-75;  
Matches 411; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 2938 TGGTCTTACTGAGCTGGGTGGTGTCTGTTTGGAGCTTATTAGAGTCTCTAGTCTTCTTA 2997  
Db 47 TGGTCTTACTGAGCTGGGTGGTGTCTGTTTGGAGCTTATTAGAGTCTCTAGTCTTCTTA 106  
QY 2998 CTTATAAAGTAGAAAATGGTGAGATTGTTTCTTTTCTACCKTAAAGGAGATGTTAAGA 3057  
Db 107 CTTATAAAGCAGAAACGGGAGATTGTTTGTGTTTCTACCTTAAAGGAGATGCAAGA 166  
QY 3058 AACAAATGAATGTCCTTTTTCAAACCTTTATTTGACAAGTGATTTTCAAGTCTGTGTTCAAAA 3117  
Db 167 AACAAATGAATGTCCTTTTTCAAACCTTTATTTGACGAGTGAATTTTCAGTCTGTGTTCAAAA 226  
QY 3118 ATATATTTCATGTACCTGTGATCCAGCAAGAGGGAGTTCCAGTCAAGAGTCACTCAACT 3177  
Db 227 ATATATTTCATGTACCTGTGATCCAGCAAGAGGGAGTTCCAGTCAAGAGTCACTCAACT 286  
QY 3178 GATTAGTTGTTTATAGAGAATGAGAAAATGGAACAGTGAAGGAATGGAGCCATATTTCCATGA 3237  
Db 287 GATTAGTTGTTTATAGAGAATGAGAAAATGGAACAGTGAAGGAATGGAGCCATATTTCCATGA 346  
QY 3238 CTTCCCTTTGTAACAGAGCAACAGAGGGACAGAGGGTGGCCTCTACATCACTCTCAC 3297  
Db 347 CTTCCCTTTGTAACAGAGCAACAGAGGGACAGAGGGTGGCCTCTACATCACTCTCAC 406  
QY 3298 CTTCCAAATCTTGTGGAAGTGCATCTACTTGGCAGAACCAATTAACCTTACCTCAAGTT 3357  
Db 407 CTTCCAAATCTTGTGGAAGTGCATCTACTTGGCAGAACCAATTAACCTTACCTCAAGTT 466  
QY 3358 CT 3359  
Db 467 CT 468

Search completed: August 30, 2005, 06:01:37  
Job time : 2165 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Query			DB	ID	Description	
	Score	Match	Length				
1	4108.4	93.3	4386	3	US-09-300-008B-1	Sequence 1, Appli	
2	552.2	12.5	646	4	US-09-401-064-343	Sequence 343, App	
C	3	84	601	4	US-09-949-016-79057	Sequence 79057, A	
	4	82.8	1.9	784019	US-09-949-016-14033	Sequence 14033, A	
5	82.8	1.9	828152	4	US-09-949-016-12077	Sequence 12777, A	
C	6	81.6	1.9	3285	US-09-573-080A-425	Sequence 425, App	
	7	79.8	1.8	7218	US-08-232-463-14	Sequence 14, Appli	
C	8	78	1.8	75216	US-09-949-016-14061	Sequence 14061, A	
	9	75.4	1.7	141454	US-09-949-016-13055	Sequence 12055, A	
10	75.4	1.7	141455	4	US-09-949-016-17190	Sequence 17190, A	
C	11	74.8	1.7	173992	4	US-09-949-016-13379	Sequence 13379, A
C	12	69.6	1.6	612	US-09-902-540-1357	Sequence 1357, Ap	
C	13	68.2	1.5	69709	4	US-09-949-016-15784	Sequence 15784, A
C	14	68.2	1.5	146039	4	US-09-949-016-12449	Sequence 12449, A
C	15	67.8	1.5	601	4	US-09-949-016-31749	Sequence 31749, A
C	16	67.8	1.5	601	4	US-09-949-016-143186	Sequence 143186,
C	17	67.8	1.5	1141	4	US-09-806-708B-22	Sequence 22, Appli
18	67.4	1.5	29357	4	US-09-949-016-16576	Sequence 16676, A	
C	19	67.2	1.5	601	4	US-09-949-016-152075	Sequence 152075,
C	20	67.2	1.5	601	4	US-09-949-016-152148	Sequence 152148,
C	21	67.2	1.5	601	4	US-09-949-016-152221	Sequence 152221,
C	22	67.2	1.5	601	4	US-09-949-016-152294	Sequence 152294,
C	23	67.2	1.5	601	4	US-09-949-016-159380	Sequence 159380,
C	24	67.2	1.5	601	4	US-09-949-016-159453	Sequence 159453,
C	25	67.2	1.5	601	4	US-09-949-016-159526	Sequence 159526,
C	26	67.2	1.5	601	4	US-09-949-016-159599	Sequence 159599,
27	67.2	1.5	78946	4	US-09-949-016-12396	Sequence 12396, A	

Db 301 GTGGGGATGGTATTTACTTTTGGAGTGTGTTGGAAGTAAATTCAGAATAGAGTATGAGCC 360  
Qy 388 TTTGGTTGCATGCTCTTCTGTTTAGATGTCTCTGGGAAAACTGCTTTAAATCAAGCTAT 447  
Db 361 TTTGGTTGCATGCTCTTCTGTTTAGATGTCTCTGGGAAAACTGCTTTAAATCAAGCTAT 420  
Qy 448 ATTGCAACTTGGAGGATTTACTGTAAACAATTGGACAGAAGAAATGCACCTCACCTTGTCAT 507  
Db 421 ATTGCAACTTGGAGGATTTACTGTAAACAATTGGACAGAAGAAATGCACCTTGTCAT 480  
Qy 508 GGTATCAGTGAAGTTTACCATTAATAACAATATGTGCACCTCATTTGTGGAAGTGTCAATGT 567  
Db 481 GGTATCAGTGAAGTTTACCATTAATAACAATATGTGCACCTCATTTGTGGAAGTGTCAATGT 540  
Qy 568 AAAGCCAGATATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAAAGCAGGCTCCACA 627  
Db 541 AAAGCCAGATATTTTACTGAAATTCCTGAAAGCAGTTTGAGTCCAAAGCAGGCTCCACA 600  
Qy 628 AATTGAAAGTTTTCACCCACTCTTGATGAACCATCTATTGGAAGTAAAAATGTTGATCT 687  
Db 601 AATTGAAAGTTTTCACCCACTCTTGATGAACCATCTATTGGAAGTAAAAATGTTGATCT 660  
Qy 688 GTCAGGACGCGAGGAAAGAAAAAATACTTCAAGGGGAAAAAATAATTTATATTTTGAATGC 747  
Db 661 GTCAGGACGCGAGGAAAGAAAAAATACTTCAAGGGGAAAAAATAATTTATATTTTGAATGC 720  
Qy 748 CAAACAGCATAGAAATTTGAGTTCGCGAGTTGCTTTGGAGTGGGGAAGCTAGGTTGAT 807  
Db 721 CAAACAGCATAGAAATTTGAGTTCGCGAGTTGCTTTGGAGTGGGGAAGCTAGGTTGAT 780  
Qy 808 AACAGAAGAGATGAAGAAGAACATAATTTCTTTTGGCTCCGGGAAACGTGTGTTGTCGA 867  
Db 781 AACAGAAGAGATGAAGAAGAACATAATTTCTTTTGGCTCCGGGAAACGTGTGTTGTCGA 840  
Qy 868 TACAGGAATAACAAACTCAGACCTTAAATTCCTGACTGTCAAGAAGAAATGGAATCAGTC 927  
Db 841 TACAGGAATAACAAACTCAGACCTTAAATTCCTGACTGTCAAGAAGAAATGGAATCAGTC 900  
Qy 928 AATAATGGATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTGAGCAGAAATTTGGATT 987  
Db 901 AATAATGGATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTGAGCAGAAATTTGGATT 960  
Qy 988 GGGGTGATTTTCATGACTACAAGAAATTAATCTGTGATCCTCAGGGCCATCCCAAGTACAGG 1047  
Db 961 GGGGTGATTTTCATGACTACAAGAAATTAATCTGTGATCCTCAGGGCCATCCCAAGTACAGG 1020  
Qy 1048 ATTAAGAACAACAACTCCAGGACCAAGCCCTTCACAAGGCGTGTCAAGTGAATAAACT 1107  
Db 1021 ATTAAGAACAACAACTCCAGGACCAAGCCCTTCACAAGGCGTGTCAAGTGAATAAACT 1080  
Qy 1108 AATGCCAAGGCCCGGAGTGAACACTACAACATACGTAGTGCACACAGATCAGAGCAAGC 1167  
Db 1081 AATGCCAAGGCCCGGAGTGAACACTACAACATACGTAGTGCACACAGATCAGAGCAAGC 1140  
Qy 1168 AGATACATGGGATTTGAGTGAAGGCAAGGCAAGAAATCAAAAGTCTCCAAAATGGAACAAA 1227  
Db 1141 AGATACATGGGATTTGAGTGAAGGCAAGGCAAGAAATCAAAAGTCTCCAAAATGGAACAAA 1200  
Qy 1228 ATTCAGAAATGCTTTCAAGACGACCCACTGCTGAAAGGAGTCTCTGCAAAACAAAGCTCTAA 1287  
Db 1201 ATTCAGAAATGCTTTCAAGACGACCCACTGCTGAAAGGAGTCTCTGCAAAACAAAGCTCTAA 1260  
Qy 1288 TAATAATAGTATGGTATCAAAATCTTTGGCTAAGATGAGAAATCCCAAACTATCAGCTTTC 1347  
Db 1261 TAATAATAGTATGGTATCAAAATCTTTGGCTAAGATGAGAAATCCCAAACTATCAGCTTTC 1320  
Qy 1348 ACCAACTAAATTCGCCAAGTATAATAAAGTAAAGATAGGGCTTCTCAGCAGCAGCAGAC 1407  
Db 1321 ACCAACTAAATTCGCCAAGTATAATAAAGTAAAGATAGGGCTTCTCAGCAGCAGCAGAC 1380  
Qy 1408 CAACTCCATCAGAAACTACTTTTCAGCCGCTTACCAAAAAAAGGGAAGGATGAAGAAA 1467  
Db 1381 CAACTCCATCAGAAACTACTTTTCAGCCGCTTACCAAAAAAAGGGAAGGATGAAGAAA 1440

Qy 1468 TCAAGAAATGCTCTTCATGCBAATCAGCAAGATAGAAAGCTCTTCTCTCTTTTAGAACCA 1527  
Db 1441 TCAAGAAATGCTCTTCATGCBAATCAGCAAGATAGAAAGCTCTTCTCTCTTTTAGAACCA 1500  
Qy 1528 AACCAAACTGCTACACCCCTCATTTGTGAAATAAAGGAGCAGCATCTATCTGAGAATGA 1587  
Db 1501 AACCAAACTGCTACACCCCTCATTTGTGAAATAAAGGAGCAGCATCTATCTGAGAATGA 1560  
Qy 1588 GCCTGTGACACAAAATCTCAGCAATAAATTATTTACAGATACAGATTTAAATCTATTGT 1647  
Db 1561 GCCTGTGACACAAAATCTCAGCAATAAATTATTTACAGATACAGATTTAAATCTATTGT 1620  
Qy 1648 GAAAAATTTCTGCCAGTAAATCTCATGTCAGAAAAAGCTTAAGATCAATATAAAAAAGGA 1707  
Db 1621 GAAAAATTTCTGCCAGTAAATCTCATGTCGAGAAAAAGCTTAAGATCAATATAAAAAAGGA 1680  
Qy 1708 AATGGATGATGTGGCCATAGAAAGATGAAGTATTTGGAACAGTTATTTCAAGGACACAAAACC 1767  
Db 1681 AATGGATGATGTGGCCATAGAAAGATGAAGTATTTGGAACAGTTATTTCAAGGACACAAAACC 1740  
Qy 1768 AGAGTTAGAAATTTGATGTGAAAGTTTCAAAAAACAGGAGGAAGATGTCAATGTTAGAAAAAG 1827  
Db 1741 AGAGTTAGAAATTTGATGTGAAAGTTTCAAAAAACAGGAGGAAGATGTCAATGTTAGAAAAAG 1800  
Qy 1828 GCCAAGGATGGATATAGAAAAAATAATGACACTTTTCAGTGATGAAGCAGTACCAGAAAGTAG 1887  
Db 1801 GCCAAGGATGGATATAGAAAAAATAATGACACTTTTCAGTGATGAAGCAGTACCAGAAAGTAG 1860  
Qy 1888 CAAAAATCTCTCAAGAAAAATGAAATTTGGGAAGAAACGTGAACCTCAAGGAAGACTCATATG 1947  
Db 1861 CAAAAATCTCTCAAGAAAAATGAAATTTGGGAAGAAACGTGAACCTCAAGGAAGACTCATATG 1920  
Qy 1948 GTCAGCTAAAGAAATATCTAACAATGACAAACTTCAGGATGATGATGAGATGCTTCCAAA 2007  
Db 1921 GTCAGCTAAAGAAATATCTAACAATGACAAACTTCAGGATGATGATGAGATGCTTCCAAA 1980  
Qy 2008 AAGCTGTTTATTTGACTGAAATTTAGATCACTGTGTGATTAATAAACTCTACTTCCAGAAATCC 2067  
Db 1981 AAGCTGTTTATTTGACTGAAATTTAGATCACTGTGTGATTAATAAACTCTACTTCCAGAAATCC 2040  
Qy 2068 GTCTGGCATAAATGATGATTTATGTCATCAATAAAAAATTTCAAGAAATTTCAAAAAGGTCA 2127  
Db 2041 ATCTGSCATAAATGATGATTTATGTCATCACTAAAAAATTTCAAGAAATTTCAAAAAGGTCA 2100  
Qy 2128 ATATCCTGAGCAGGAAAACTTCCACATCATTTGAGGATCAGATCTAATAGCTCATCA 2187  
Db 2101 ATATCCTGAGCAGGAAAACTTCCACATCATTTGAGGATCAGATCTAATAGCTCATCA 2160  
Qy 2188 TGCTCGAAAGAAATACAGAACTAGAAAGTGGCTAAGCGAGGAAATGGAGGTACAAAATCA 2247  
Db 2161 TGCTCGAAAGAAATACAGAACTAGAAAGTGGCTAAGCGAGGAAATGGAGGTACAAAATCA 2220  
Qy 2248 ACATGCAAAAAGAGAGTCTCTTGTGATGATCTTTTATAGATCAATCTCTTTTAAAAAG 2307  
Db 2221 ACATGCAAAAAGAGAGTCTCTTGTGATGATCTTTTATAGATCAATCTCTTTTAAAAAG 2280  
Qy 2308 GAGAGATAAATGAGGATTTTAAAAAGAGCCATCGAAAAAATCTTCTAGTAGCATCTAC 2367  
Db 2281 GAGAGATAAATGAGGATTTTAAAAAGAGCCATCGAAAAAATCTTCTAGTAGCATCTAC 2340  
Qy 2368 TTCAGGCCAACAGGTTATATGAATATATAGTATAGAGCGATTTAAGTTTCAATGTT 2427  
Db 2341 TTCAGGCCAACAGGTTATATGAATATATAGTATAGAGCGATTTAAGTTTCAATGTT 2400  
Qy 2428 TTTATGGCCCTAAATTTTATTAATTAATAATGCAAAAACTTTGATCTTTTGTATGTAACAAAT 2487  
Db 2401 TTTATGGCCCTAAATTTTATTAATTAATAATGCAAAAACTTTGATCTTTTGTATGTAACAAAT 2460  
Qy 2488 TGTGTTGTCGTTTTTCAGGCTTTGTCATGTCATTTTTTTCATTTTTTAAATGTCGTTTG 2547  
Db 2461 TGTGTTGTCGTTTTTCAGGCTTTGTCATGTCATTTTTTTCATTTTTTAAATGTCGTTTG 2520

Qy	2548	TTTATTAATAAGTAAATATATAGTCAAGTTCAAAATTCCTAAATRTACCTAAGGTAAAGGAC	2607
Db	2521	TTTATTAATAAGTAAATATATAGTCAAGTTCAAAATTCCTAAATRTACCTAAGGTAAAGGAC	2579
Qy	2608	TAAAGTCACCTTCACCAATGTCTAGCTACT-----	2640
Db	2580	TAAAGTCACCTTCACCAATGTCTAGCTACT-----	2639
Qy	2641	-----	2648
Db	2640	TACTCATTTCTTATGAATCTTTCCAGGAGTTTTTGAGTCCTATTTCAAATTCCTATTTTAA	2699
Qy	2649	AATAATTTCTTACACAAATGANTAGCATAAACATATGCAAGTGTCTTACACCTTGCTTTTTTAA	2708
Db	2700	AATAATTTCTTACACAAATGANTAGCATAAACATATGCAAGTGTCTTACACCTTGCTTTTTTAA	2759
Qy	2709	CTTAGTAGAATTTAAAAATTTATAGGAATATCAATATATATGTTTTTAAATATTTTCTTTTTC	2768
Db	2760	CTTAGT-AGATTTAAAAATTTATAGGAATATCAATATATATGTTTTTAAATATTTTCTTTTTC	2818
Qy	2769	CATTATGCTGTAGTCTTACCTTAAACTCTGGTGATCCAAAACAAAATGGCTTCAGTGGTGCA	2828
Db	2819	CATTATGCTGTAGTCTTACCTTAAACTCTGGTGATCCAAAACAAAATGGCTTCAGTGGTGCA	2878
Qy	2829	GATGTCACTACATATGTTATTTCTAGTACTAGAACTGAAGACCATATGGAGACTTCATCAA	2888
Db	2879	GATGTCACTACATATGTTATTTCTAGTACTAGAACTGAAGACCATATGGAGACTTCATCAA	2938
Qy	2889	ACATGGGTTTAGTTTTTCCACAGAAATGAAAGACCTGTACCCCTTTTTTGGTGCTTACTCG	2948
Db	2939	ACATGGGTTTAGTTTTTCCACAGAAATGAAAGACCTGTACCCCTTTTTTGGTGCTTACTCG	2998
Qy	2949	AGCTGGGTGGGTGTCGTGTTTGAGCTTATTTTAGAGTCTTAGTTTTCCTACTTATAAAGTA	3008
Db	2999	AGCTGGGTGGGTGTCGTGTTTGAGCTTATTTTAGAGTCTTAGTTTTCCTACTTATAAAGTA	3058
Qy	3009	GAAATGTTGAGATTTGTTTTCTTTTTCTACCKTAAAGGGAGATGGTAAAGAACCAATGAATG	3068
Db	3059	GAAATGTTGAGATTTGTTTTCTTTTTCTACCKTAAAGGGAGATGGTAAAGAACCAATGAATG	3118
Qy	3069	TCTTTTTTCAAACTTTATTCACAAGTCAATTTTCAAGTCTGTGTTTCAAAAATATATTCATG	3128
Db	3119	TCTTTTTTCAAACTTTATTCACAAGTCAATTTTCAAGTCTGTGTTTCAAAAATATATTCATG	3178
Qy	3129	TACCTGTGATCCAGCAAGAGGGAGTTCAGTCAAGAGTCACTACAACTGATTAGTTGTT	3188
Db	3179	TACCTGTGATCCAGCAAGAGGGAGTTCAGTCAAGAGTCACTACAACTGATTAGTTGTT	3238
Qy	3189	TAGAGAATAGAAATGGAAACAGTGAAGAAATGGAGGCCAATATTTCCATGACTTCCCTTGTA	3248
Db	3239	TAGAGAATAGAAATGGAAACAGTGAAGAAATGGAGGCCAATATTTCCATGACTTCCCTTGTA	3298
Qy	3249	AACAGAGCAACAGAAAGGGACAGAGGCTGGCCTCTACATCACTCTCACCTTCCAAATCT	3308
Db	3299	AACAGAGCAACAGAAAGGGACAGAGGCTGGCCTCTACATCACTCTCACCTTCCAAATCT	3358
Qy	3309	TGTGGAAAGTGATCTACTCTGCCAGAACCAAAATTAACCTTCTTCCAAAGTCTTGCTGCTTG	3368
Db	3359	TGTGGAAAGTGATCTACTCTGCCAGAACCAAAATTAACCTTCTTCCAAAGTCTTGCTGCTTG	3418
Qy	3369	CAGGTGGAACCTCCAGCTGCAAGGAGTTAGGGAAATGAAGGTCTTTTTTTAAAGCTTCT	3428
Db	3419	CAGGTGGAACCTCCAGCTGCAAGGAGTTAGGGAAATGAAGGTCTTTTTTTAAAGCTTCT	3478
Qy	3429	CAGCCTTCCTAGGGACAGAAATTTGGGTGAGCCATCTGCAATTTTCTACTACAGGCATTG	3488
Db	3479	CAGCCTTCCTAGGGACAGAAATTTGGGTGAGCCATCTGCAATTTTCTACTACAGGCATTG	3538
Qy	3489	AGACCGATTAGATTTATTTGAAATATTTATAGAGAGTTATGAACACTTTAAATTTATGATGCG	3548
Db	3539	AGACCGATTAGATTTATTTGAAATATTTATAGAGAGTTATGAACACTTTAAATTTATGATGCG	3598
Qy	3549	TATGACATTTGGATAGAACATGGGATACCTTTTGAAGTAGAATTTGACAGGGCATATTAGTTG	3608

Db	3599	TATGACATTGGATAGAACATGGGATACCTTTAGAAAGTAGAATTCACAGGGCATATTAGTTG	3558
Qy	3609	ATGAAATGGAGTCATTTGAGTCCTTTAAATAGCCATGTATCATAAATACCAAGTGAAGCTG	3668
Db	3659	ATGAAATGGAGTCATTTGAGTCCTTTAAATAGCCATGTATCATAAATACCAAGTGAAGCTG	3718
Qy	3669	GTGGAACATATGCTCTCCATTTTACAGTTTAAGGAATATAATGGACAGATTAATATTGTTT	3728
Db	3719	GTGGAACATATGCTCTCCATTTTACAGTTTAAGGAATATAATGGACAGATTAATATTGTTT	3778
Qy	3729	TCGTGATGCCCCAATCCCTTTCTAAGGAAGACTGCCCTACTATAGCAGTTTTTATATT	3788
Db	3779	TCGTGATGCCCCAATCCCTTTCTAAGGAAGACTGCCCTACTATAGCAGTTTTTATATT	3838
Qy	3789	TGTCAAATTTATGAATATAATGAATGAGGAGTCTGTGTACCTCTGCTCTTTTACAAATATTG	3848
Db	3839	TGTCAAATTTATGAATATAATGAATGAGGAGTCTGTGTACCTCTGCTCTTTTACAAATATTG	3897
Qy	3849	GGTGTGTGCAGTATTTTTCCCTTTTTTAACCMTTCCCAATTCGGGTGTGTAGGTGGATGT	3908
Db	3898	G----TGTTGTCAGTATTTTCTCTTTTAACCAATCCAAATCGGTGTGTAGTGATG----	3948
Qy	3909	TTCCATTTGGGTTTTTAATTTGTATATATCCCTGATAGCTATAATTTGGGTCTATAGAAATTCCT	3968
Db	3949	TTTCAATTTGGTTTTAATTTGTATATATCCCTGATAGCTATAATTTGGGTCTATAGAAATTCCT	4008
Qy	3969	TATACATCTCAGATGCAAGTCTCTGTCGGATATACGTATTGAGATATATACACCTAGTCT	4028
Db	4009	TATACATCTCAGATGCAAGTCTCTGTCGGATATATGTATTGAGATATATACACCTAGTCT	4068
Qy	4029	GTGGCTTGACTGTTTTCTTTTATGTCCTTTTGTAGCAATAGAAATTTAAATTTTGACAAAGT	4088
Db	4069	GTGGCTTGACTGTTTTCTTTTATGTCCTTTTGTAGCAATAGAAATTTAAATTTTGACAAAGT	4128
Qy	4089	CAAAATTTATTTTTCTTTTGTGTATATTTTTCTCTCCAATTTAACCCCAAGATTTCA	4148
Db	4129	CAAAATTTATTTTTCTTTTGTGTATATTTTTCTCTCCAATTTAACCCCAAGATTTCA	4188
Qy	4149	GATATTCGTCTATTTATATATAAACCCTTATATTTTTATATTTGTGATCTACCTTGAATTGA	4208
Db	4189	GATATTCGTCTATTTATATAAACCTTTATATTTTTATATTTGTGATCTACCTTGAATTGA	4248
Qy	4209	TATGTATGTTGAAATTAGGATCAGGGTCTTTTTTTTCCCCCATACAAGTATCCAGTCA	4268
Db	4249	TATGTATGTTGAAATTAGGATCAGGGTCTTTTTTTTCCCCCATACAAGTATCCAGTCA	4308
Qy	4269	TTGTAAACACTGTTTTATTGAAAGAATTTATCTTTTCCTCATTTAAATACCTTTGCCAATTAGT	4328
Db	4309	TTGTAAACACTGTTTTATTGAAAGAATTTATCTTTTCCTCATTTAAATACCTTTGCCAATTAGT	4368
Qy	4329	AAAAAATCAATTAACCAT	4346
Db	4369	AAAAAATCAATTAACCAT	4386

## RESULT 2

US-09-401-064-343/C  
; Sequence 343, Application US/09401064  
; Patent No. 6623923  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Meagher, Madeline Joy  
; APPLICANT: Stolk, John A.  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE  
; FILE REFERENCE: 210121.471C2  
; CURRENT APPLICATION NUMBER: US/09/401,064  
; CURRENT FILING DATE: 1999-09-22



US-09-949-016-14033

Query Match 1.9%; Score 82.8; DB 4; Length 784019;  
Best Local Similarity 56.6%; Pred. No. 3.6e-09;  
Matches 248; Conservative 2; Mismatches 154; Indels 34; Gaps 4;  
QY 3945 TATAATGGGTCATAGAAATCTTTATACATTTCTAGATGCAAGTCTCTTGYCGGATATAC 4004  
DB 467327 TAATATTGAATTATAGAAATCTTATATATATTTTGGATACATCTTTTTTTTCATATGT 467386  
QY 4005 GTATTGAGATATTACACCTAGTCTGGCTTGACTGTTTCTTATGCTCTTTTGATGAAT 4064  
DB 467387 ACTGTGA-ATTTTCTCTTGTCTGTCTATTTTCCCTTAATGGTATCTTTTGAAGAAGC 467445  
QY 4065 AGAAGTTTTAAATTTTGCAAGGTCAAATTTATTTTTTTTCTTTTGTGATATTTTTTCT 4124  
DB 467446 AGAAGTTATAAATACTGATAGTGTCCAAATTAATCTTTTTCATTTGTTAGTGCTTTCTTG 467505  
QY 4125 CTCC-----AATTAACCCCAAGATTTCCAGATATT 4154  
DB 467506 TTCTAAGAACCTTTTACCAATCACCAGGTCAAAAAATATTTTCTATGTTTCTTCTAGA 467565  
QY 4155 CTGCTCTATTATATAAATTTTATATTTTATATTTG--ATCTACCTTGAATTTGATG 4212  
DB 467566 CACTCATAACTTCAATCTTTTACATTCAGATCTATGTTCCATCTCAAAATTAATTTTGTGA 467625  
QY 4213 TATGTTGTGAATTTATGGATCAGGTTCTTTTTTCCCCATACAAAGTATCCAGTCAATGT 4272  
DB 467626 TATGGTGTGAGGTAATGGTAGAGTTTATTTTCTCTCCATAGACTCTCCAATTTGTCT 467685  
QY 4273 AACACTGTTTATGAAGAATTAATCTTCTCTCATTAATAATTAATCTTGCCA-ATTAGTAA 4331  
DB 467686 AGCACTATTGTTTCAAGACTAATCTCGGCTATTCAATTTCTTGGCATCTTTGTCAA 467745  
QY 4332 AATCAATTAACCATRMA 4349  
DB 467746 AATCAGTTGACTATATA 467763

RESULT 5

US-09-949-016-12777  
; Sequence 12777, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12777  
; LENGTH: 828152  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(828152)  
; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-12777

Query Match 1.9%; Score 82.8; DB 4; Length 828152;  
Best Local Similarity 56.6%; Pred. No. 3.7e-09;  
Matches 248; Conservative 2; Mismatches 154; Indels 34; Gaps 4;  
QY 3945 TATAATGGGTCATAGAAATCTTTATACATTTCTAGATGCAAGTCTCTTGYCGGATATAC 4004

DB 463460 TAATATTGAATTATAGAAATCTTATATATATTTTGGATACATTTCTTTTTTCATATGT 463519  
QY 4005 GTATTGAGATATTACACCTAGTCTGGCTTGACTGTTTCTTATGCTTTTGTGATGAAT 4064  
DB 463520 ACTGTGA-ATTTTCTCTTGTCTGTCTATTTTCCCTTAATGGTATCTTTTGAAGAAGC 463578  
QY 4065 AGAAGTTTTAAATTTTGCAAGGTCAAATTTATTTTTTTTCTTTTGTGATATTTTTTCT 4124  
DB 463579 AGAAGTTATAAATACTGATAGTGTCCAAATTAATCTTTTTCATTTGTTAGTGCTTTCTTG 463638  
QY 4125 CTCC-----AATTAACCCCAAGATTTCCAGATATT 4154  
DB 463639 TTCTAAGAACCTTTTACCAATCACCAGGTCAAAAAATATTTTCTATGTTTCTTCTAGA 463698  
QY 4155 CTGCTCTATTATATAAATTTTATATTTTATATTTG--ATCTACCTTGAATTTGATG 4212  
DB 463699 CACTCATAACTTCAATCTTTTACATTCAGATCTATGTTCCATCTCAAAATTAATTTTGTGA 463758  
QY 4213 TATGTTGTGAATTTATGGATCAGGTTCTTTTTTCCCCATACAAAGTATCCAGTCAATGT 4272  
DB 463759 TATGGTGTGAGGTAATGGTAGAGTTTATTTTCTCTCCATAGACTCTCCAATTTGTCT 463818  
QY 4273 AACACTGTTTATGAAGAATTAATCTTCTCTCATTAATAATTAATCTTGCCA-ATTAGTAA 4331  
DB 463819 AGCACTATTGTTTCAAGACTAATCTCGGCTATTCAATTTCTTGGCATCTTTGTCAA 463878  
QY 4332 AATCAATTAACCATRMA 4349  
DB 463879 AATCAGTTGACTATATA 463896

RESULT 6

US-09-573-080A-425/c  
; Sequence 425, Application US/09573080A  
; Patent No. 6828097  
; GENERAL INFORMATION:  
; APPLICANT: JOAN, KNOLL  
; APPLICANT: ROGAN, PETER  
; TITLE OF INVENTION: SINGLE COPY GENOMIC HYBRIDIZATION PROBES AND METHOD OF GENERAT  
; FILE REFERENCE: 30307  
; CURRENT APPLICATION NUMBER: US/09/573,080A  
; CURRENT FILING DATE: 2000-05-16  
; NUMBER OF SEQ ID NOS: 479  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 425  
; LENGTH: 3285  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: repeat region  
; LOCATION: (1)...(3285)  
; OTHER INFORMATION: lme\_orf2  
; NAME/KEY: misc feature  
; OTHER INFORMATION: n is a, c, g or t  
; PUBLICATION INFORMATION:  
; PUBLICATION INFORMATION:  
; AUTHORS: Jurka, J; Walichewicz, J; Milosavljevic, A  
; TITLE: Prototypic sequences for human repetitive DNA  
; JOURNAL: Journal of Molecular Evolution  
; VOLUME: 35  
; ISSUE: 4  
; PAGES: 286-291  
; DATE: 1992-10-  
; DATABASE ACCESSION NUMBER: Database of repetitive elements (repbase)  
; DATABASE ENTRY DATE: 1996-01-26  
; DATABASE ENTRY DATE: 1996-01-26

US-09-573-080A-425

Query Match 1.9%; Score 81.6; DB 4; Length 3285;  
Best Local Similarity 58.8%; Pred. No. 3.2e-10;  
Matches 284; Conservative 7; Mismatches 137; Indels 55; Gaps 7;

Qy	3920	TTTTAAATTTGTAATACCTCGATAGCTATAAATTTGGGTCAATGAAATCTTTATACATTCCTA	3979
Db	3203	TTTTTAAATTTGGGTTAATTTGTCTTTTATATTGAGTTTGAAGAGTCTTTTATATATTCTG	3144
Qy	3980	GATCGAAGTCTCTTGYCGGATATACGTAATTTGAGATA--TTACACCTAGTCTGTGGCTTGA	4037
Db	3143	GATACAAGTCTTTGTGCAGATATATGTTTTGCAATAATTTTCTCCAGTCTGTGGCTTGT	3084
Qy	4038	CTGTGTTTCTTTA-----TGTCTTTTGTGAATPAGAAGTTTTAAATTTTGTGACAAGTGCA	4090
Db	3083	CTTTTCATTTCTTGATRGTCCTTTTGAAGAGCAGAAGTTTTTAAATTTTGTATGAAGTCC	3024
Qy	4091	AAATTTAT-----TTTTTCTTTTGTGTAATATTTTTTCTCTCCAAT	4131
Db	3023	AAATTTATCAATTTTCTTTTGTGTCCTTTTGTGTCATCTAAGAATCTGTTG	2964
Qy	4132	TTAACCCCAAGATTTTCAGATAT-----TCTGCTCTATTATATAAATCTTTATATTTTTA	4184
Db	2963	CCTARCCCAAGGTCACGAAGATTTTCTCTATGTTTCTTCTAARAGTTTATATAGTTTTA	2904
Qy	4185	TATTTGTGATCTACCT-----TGAATTGATGATGATGTTGTGCAATAT	4227
Db	2903	GNTTTTACATTTAGGCTCTATATCCATTTTGTGAGTTAATTTTGTATATGTTGTGAGTAG	2844
Qy	4228	GGATCAGGTTCTTTTTTCCCCCATACAAGTATCCAGTCAATTTGTAACACTGTTTATTTGA	4287
Db	2843	GGGTCKAATTCATTTTTT--TGCAATATGGATATCCAGTTGTTCCAGACCAATTTGTTGA	2786
Qy	4288	AAGAATTTATCCTTTCTCATTTAAATTTACCTTTGCCA-ATTAGTAAAAAATCAATTAACCAT	4346
Db	2785	AAAGACTATTCTTTTCCCCCATTTGAATTCCTTTGGCACCTTTTGTCAAAAAATTATTGGCCAT	2726
Qy	4347	RMA	4349
Db	2725	ATA	2723

## RESULT 7

```

US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:

```

RESULTS

```

US-09-949-016-14061/c
; Sequence 14061, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14061
; LENGTH: 75216
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(75216)
; OTHER INFORMATION: n = A,T,C or G

```



```
QY 4177 TATTTTATATTTGATCTACCTGAAATTGAT--ATGTATGTTGTAATTTATGATCAG 4234
Db 126603 TTTAGCCCTATAGTCATCTTTCTTTTTCATGAATGGTGTAGTGAATTCGA 126662
QY 4235 GGTCTTTTTTCCCCATACAGTATCCAGTCATTTGTAACACTGTTTATTTGAAAGAAATT 4294
Db 126663 AGTTCAGTTTCTTTTCATACCGATATCCAGTTGTTCCAGCACAAATTAATTGMAAGATT 126722
QY 4295 ATCTTTTCTTCATTAATTAACCTTG 4319
Db 126723 TTTACTTTCTCATTAATGTTTCTCTG 126747
```

```
RESULT 11
US-09-949-016-13379/c
; Sequence 13379, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13379
; LENGTH: 173992
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13379
```

```
Query Match 1.7%; Score 74.8; DB 4; Length 173992;
Best Local Similarity 57.3%; Pred. No. 1.7e-07;
Matches 172; Conservative 1; Mismatches 123; Indels 4; Gaps 2;

QY 3839 ACAAAATATTTGGGTGTGTCAGATATTTTCCCTTTTAAACCMTTCGCCAATTCGGGTGTGT 3898
Db 144215 ACTAAGGTGTGTGATAGTCTTTCAITGTGTAATTTGCCACCTACACATCTTCTCGATG 144156
QY 3899 AGGTGATGTTCCATTTGGGTTTTTAATTTGTATATCCCTGATAGCTATAAATTTGGGTGAT 3958
Db 144155 AACTGTCTGTTCAAATTTTGGCCATTTTAAATGCAGTTGTTGTTTCTTGAGTTT 144096
QY 3959 AGNAATCTTTATACATTTAGATGCAAGTCTCTTGYCGGATATACGTATTGAGATATTA 4018
Db 144095 AAGAATCTTTATACATTTGGATACAGTCTCTCAGACACATATGATTTAAATTTTA 144036
QY 4019 CACTAGTCTGTGGCTTGACTGT---TTTCTTTATGCTTTTGTGATGAATGAAGTTTAA 4075
Db 144035 -TCCAGTCTGTGGCTGTGCTTTTCATTTCTTAACAGTTTGAAGACGAGAAATACCTA 143977
QY 4076 ATTTGACAGGTCAAATTTATTTTTTCTTTTGTGATATTTTTTCTCTCAATTTAA 4135
Db 143976 ATTTGATGAAGTCCAATTTATTAATGTTTATGGATTGCACTTTTTGTGCATATCTAA 143917
```

```
RESULT 12
US-09-902-540-1357/c
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
```

```
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(612)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match 1.6%; Score 69.6; DB 4; Length 612;
Best Local Similarity 47.9%; Pred. No. 1.4e-07;
Matches 224; Conservative 1; Mismatches 242; Indels 1; Gaps 1;

QY 3779 TTTTATATTTGTCATATTAATGAATGAGGAGTCTCGTACCTCCTCTCTTT 3838
Db 504 TTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 445
QY 3839 ACAAAATATTTGGGTGTGTCAGATATTTTCCCTTTTAAACCMTTCGCCAATTCGGGTGTGT 3898
Db 444 TTTAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 385
QY 3899 AGGTGATGTTTCCATTTGGGTTTTTAATTTGTATATCCCTGATAGCTATAAATTTGGGTGAT 3958
Db 384 ATTTTATTTTTTTTATTTT-TTTTTTATTTATTTTTTTTTTTTTTANTTTTTTTT 326
QY 3959 AGAAATCTTTATACATTTAGATGCAAGTCTCTTGYCGGATATACGTATTGAGATATTA 4018
Db 325 TTTTATTTTATTTTTTTTTTTTTTTTTTTTAAATTTTTTTTTTTTATTTTTTTTATTTT 266
QY 4019 CACTAGTCTGTGGCTTGACTGTTTCTTTATGCTTTTGTGATGAATGAAGTTTAAATTT 4078
Db 265 ATTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 206
QY 4079 TTGACAAAGTCAAATTTATTTTTTCTTTTGTGATATTTTCTCTCCAATTTAACCC 4138
Db 205 TTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 146
QY 4139 CAAGATTTTCAGATTTCTGCTCTATTATATAAACTTTATATTTTATATTTTCTGATCTAC 4198
Db 145 TTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 86
QY 4199 CTGTAATTCATATGATGTTGTTGAATTTATGATCAGGTTCTTTTTTT 4246
Db 85 TTTTATTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 38
```

```
RESULT 13
US-09-949-016-15784/c
; Sequence 15784, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15784
```



Db 126 A 126

Search completed: August 30, 2005, 15:39:18  
Job time : 684 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 05:17:43 ; Search time 2560 Seconds  
(without alignments)  
11253.603 Million cell updates/sec

Title: US-09-837-138-1

Perfect score: 4403

Sequence: 1 ttccgcagcagcgcggttg.....accgcggtagctcagct 4403

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7331713 seqs, 3271544945 residues

Total number of hits satisfying chosen parameters: 14663426

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_NA.\*  
1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/prodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/prodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/prodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/prodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/prodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/prodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/prodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/prodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/prodata/2/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/prodata/2/pubpna/US10G\_PUBCOMB.seq.\*  
20: /cgn2\_6/prodata/2/pubpna/US10H\_PUBCOMB.seq.\*  
21: /cgn2\_6/prodata/2/pubpna/US10I\_PUBCOMB.seq.\*  
22: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
23: /cgn2\_6/prodata/2/pubpna/US11A\_PUBCOMB.seq.\*  
24: /cgn2\_6/prodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
26: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	564.6	12.8	622	19 US-10-474-495-34	Sequence 34, Appl
C 2	552.2	12.5	646	9 US-09-922-217-343	Sequence 343, App
C 3	552.2	12.5	646	9 US-09-813-263-343	Sequence 343, App
C 4	552.2	12.5	646	13 US-10-025-380-343	Sequence 343, App
C 5	544.8	12.4	561	11 US-09-969-034-1463	Sequence 1463, Ap
6	537.4	12.2	541	9 US-09-878-178-1737	Sequence 1737, Ap
7	537.4	12.2	541	13 US-10-046-935-1737	Sequence 1737, Ap

8	537.4	12.2	541	14	US-10-146-502-1737	Sequence 1737, Ap
9	507	11.5	587	15	US-10-102-524-532	Sequence 532, App
10	478.4	10.9	481	17	US-10-242-535A-26197	Sequence 26197, A
11	478.4	10.9	481	18	US-10-085-783A-26197	Sequence 26197, A
12	438.2	10.0	439	17	US-10-242-535A-7179	Sequence 7179, Ap
13	438.2	10.0	439	18	US-10-085-783A-7179	Sequence 7179, Ap
14	405.6	9.2	468	10	US-09-918-995-13834	Sequence 13834, A
C 15	362	8.2	483	11	US-09-969-034-1974	Sequence 1974, A
16	258.4	5.5	360	9	US-09-783-590-3449	Sequence 3449, Ap
17	240.8	5.5	267	20	US-10-425-115-120439	Sequence 120439,
18	239.4	5.4	272	9	US-09-923-876-5167	Sequence 5167, Ap
19	239.4	5.4	272	10	US-09-923-876-5167	Sequence 5167, Ap
C 20	212.6	4.8	475	9	US-09-917-800A-803	Sequence 803, App
21	191	4.3	404	9	US-09-783-590-8929	Sequence 8929, Ap
22	167.6	3.8	315	9	US-09-960-352-8727	Sequence 8727, Ap
23	121.8	2.8	540	21	US-10-867-578-1	Sequence 1, Appli
24	119.4	2.7	442	9	US-09-864-761-11287	Sequence 11287, A
25	114	2.6	114	9	US-09-864-761-27987	Sequence 27987, A
26	97.6	2.2	6237	15	US-10-311-455-331	Sequence 331, App
C 27	92.4	2.1	6237	15	US-10-311-455-332	Sequence 332, App
C 28	81.6	1.9	3285	10	US-09-854-857-425	Sequence 425, App
C 29	81.6	1.9	3285	21	US-10-786-970A-425	Sequence 425, App
30	80	1.8	1300	19	US-10-668-749A-1	Sequence 1, Appli
31	74.6	1.7	1980090	20	US-10-719-393-6815	Sequence 6815, Ap
32	74.6	1.7	1980090	21	US-10-741-600-17676	Sequence 17676, A
C 33	73.6	1.7	539	14	US-10-198-846-1369	Sequence 1369, Ap
34	72.2	1.6	7442	18	US-10-221-714A-409	Sequence 409, App
35	72	1.6	9539	14	US-10-239-676-52	Sequence 52, Appl
36	72	1.6	9539	15	US-10-240-453-54	Sequence 54, Appl
C 37	71.4	1.6	739	20	US-10-425-115-55020	Sequence 55020, A
38	71.4	1.6	1062	20	US-10-425-115-120013	Sequence 120013,
C 39	71.2	1.6	322885	21	US-10-679-366-3	Sequence 3, Appli
40	71	1.6	14006	15	US-10-311-455-1931	Sequence 1931, Ap
C 41	70	1.6	6056	15	US-10-311-455-753	Sequence 753, App
C 42	70	1.6	6056	18	US-10-221-714A-85	Sequence 85, Appl
43	69.8	1.6	1243	20	US-10-425-115-172717	Sequence 172717,
44	69.6	1.6	15548	15	US-10-311-455-2128	Sequence 2128, Ap
45	69.6	1.6	3673778	16	US-10-312-841-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-10-474-495-34  
; Sequence 34, Application US/10474495  
; Publication No. US20040171003A1  
; GENERAL INFORMATION:  
; APPLICANT: YOSHIKAWA, Yoshie et al.  
; TITLE OF INVENTION: CANCER-ASSOCIATED GENES  
; FILE REFERENCE: 1422-0606P  
; CURRENT APPLICATION NUMBER: US/10474,495  
; PRIOR FILING DATE: 2003-10-09  
; PRIOR APPLICATION NUMBER: JP 2001-112039  
; PRIOR FILING DATE: 2001-04-10  
; PRIOR APPLICATION NUMBER: JP 2001-290193  
; PRIOR FILING DATE: 2001-09-21  
; NUMBER OF SEQ IDS: 264  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 34  
; LENGTH: 622  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: variation  
; LOCATION: 553, 597  
; OTHER INFORMATION: n is a or c or g or t.  
US-10-474-495-34

Query Match 12.8%; Score 564.6; DB 19; Length 622;  
Best Local Similarity 97.6%; Pred. No. 1.1e-113;  
Matches 604; Conservative 0; Mismatches 11; Indels 4; Gaps 3;

QY 840 TTTTGGCTCGGGAACGTGTGTTGATACAGGATACAACTCACAGACTTAATTC 899  
Db |||||||  
1 TTTTGGCTCGGGAACGTGTGTTGATACAGGATACAACTCACAGACTTAATTC 60  
QY 900 CTGACTGTGAGAGAAATGGATTCACTCAATATGATATGCTCCAAAGCGAGGTCTTA 959  
Db |||||||  
61 CTGACTGTGAGAGAAATGGATTCACTCAATATGATATGCTCCAAAGCGAGGTCTTA 120  
QY 960 GACCTATTCCTGAAGCAGAAATGGATTGCGGTGATTTTCATGACTACAAAGAAATTA 1019  
Db |||||||  
121 GACCTATTCCTGAAGCAGAAATGGATTGCGGTGATTTTCATGACTACAAAGAAATTA 180  
QY 1020 GTGATCCTCAGGCGCATCCAGTACAGGATTAAGACACAACTCCAGGACCAAGCCTTT 1079  
Db |||||||  
181 GTGATCCTCAGGCGCATCCAGTACAGGATTAAGACACAACTCCAGGACCAAGCCTTT 240  
QY 1080 CACAAGCGGTGTCAGTTGATGAAAACTAATGCCAAGCGCCCGAGTGAACACTACAACAT 1139  
Db |||||||  
241 CACAAGCGGTGTCAGTTGATGAAAACTAATGCCAAGCGCCCGAGTGAACACTACAACAT 300  
QY 1140 AGCTAGCTGACACAGAAATCAGAGCAAGCAGATACATGGGATTTGAGTGAAGCGCAAAAG 1199  
Db |||||||  
301 AGCTAGCTGACACAGAAATCAGAGCAAGCAGATACATGGGATTTGAGTGAAGCGCAAAAG 360  
QY 1200 AATCAAGCTCTCCAAATGGAACAAATTCAGATGCTTCAAGAGCGCACCCACTG 1259  
Db |||||||  
361 AATCAAGCTCTCCAAATGGAACAAATTCAGATGCTTCAAGAGCGCGCCACTG 420  
QY 1260 TAAAGGAGTCTCGCAAAACAAAGCTCTAATAATAGTATGATGATCAAACTACTTTGGCTTA 1319  
Db |||||||  
421 TAAAGGAGTCTCGCAAAACAAAGCTCTAATAATAGTATGATGATCAAACTACTTTGGCTTA 480  
QY 1320 AGATGAGATCCCAAACTATAGCTTTCCCAACTAAATGGCAAGTATATAAATAAGTA 1379  
Db |||||||  
481 AGATGAGATCCCAAACTATAGCTTTCCCAACTAAATGGCAAGTATATAAATAAGTA 540  
QY 1380 AGATAGGCTTC-CACAGCAGCAGCAGACCAAC-TCATCAGAACTACTTT--CAGCG 1435  
Db |||||||  
541 AGATAGGCTTTTTCAGCAGCAGCAGACCAACTCCATCAGAACTACTTTTCAGNCGG 600  
QY 1436 TCTACCAAAAAAGGAAA 1454  
Db |||||||  
601 TCTACCGGAAAAAAA 619

## RESULT 2

US-09-922-217-343/c  
; Sequence 343, Application US/09922217  
; Patent No. US200200764141

## GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Meagher, Madeleine Joy  
; APPLICANT: Stolk, John A.  
; APPLICANT: Wang, Tongcong  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Smith, Carole Lynn  
; APPLICANT: King, Gordon E.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Clapper, Jonathan D.

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS  
; FILE REFERENCE: 210121.471C13

; CURRENT APPLICATION NUMBER: US/09/922,217

; CURRENT FILING DATE: 2001-08-03

; NUMBER OF SEQ ID NOS: 1124

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 343

; LENGTH: 646

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-922-217-343

Query Match 12.5%; Score 552.2; DB 9; Length 646;

Best Local Similarity 93.0%; Pred. No. 5.8e-111;

Matches 609; Conservative 4; Mismatches 32; Indels 10; Gaps 3;

QY 3419 AAAAGCTTCTCAGCCTTCCTAGGGAACAGAAATTTGGGTGAGCCATCTGCAATTTCTACT 3478

Db |||||||

645 AAAAGCTTCTCAGCCTTCCTAGGGAACAGAAATTTGGGTGAGCCATCTGCAATTTCTACT 586

QY 3479 ACAGGCATTTGAGACCAAGTTAGATTATTGAAATATTATAGAGAGTTATGAAACACTTAAATTT 3538

Db |||||||

585 ACAGGCATTTGAGACCAAGTTAGATTATTGAAATATTATAGAGAGTTATGAAACACTTAAATTT 526

QY 3539 ATGATAGTGGTATGACATTTGGATAGAACATGGGATACCTTTAGAGTGAATTTGACAGGCG 3598

Db |||||||

525 ATGATAGTGGTATGACATTTGGATAGAACATGGGATACCTTTAGAGTGAATTTGACAGGCG 466

QY 3599 ATATTAGTTGATGAAATGGAGTCAATTTGAGTCTTTAATAGCCATGTATCAATATTACCA 3658

Db |||||||

465 ATATTAGTTGATGAAATGGAGTCAATTTGAGTCTTTAATAGCCATGTATCAATATTACCA 406

QY 3659 AGTGAAGCTGGTGAACATATGGTCTCCATTTTACAGTTAAGGAATATAATGGACAGATT 3718

Db |||||||

405 AGTGAAGCTGGTGAACATATGGTCTCCATTTTACAGTTAAGGAATATAATGGACAGATT 346

QY 3719 AATATTGTTTCTGTGATGCCCAATCCCTTTTCTAAGGAAGACTGCCCTACTATAGCAG 3778

Db |||||||

345 AATATTGTTTCTGTGATGCCCAATCCCTTTTCTAAGGAAGACTGCCCTACTATAGCAG 286

QY 3779 TTTTATATTTTGTCAATTTATGAATATAATGAATGAGGAGTTCTGGTACCTCCTGCTCTTT 3838

Db |||||||

285 TTTTATATTTTGTCAATTTATGAATATAATGAATGAGGAGTTCTGGTACCTCCTGCTCTTT 227

QY 3839 ACAATAATGGGTGTTGTCCAGTATTTTCCCTTTTAAACMTTCCCAATTCGGGTGTGT 3898

Db |||||||

226 ACAATAATGG---TGTTGTGAGTATTTTCTTTTAAACCATTTCCCAATTCGGGTGTGTAG 171

QY 3899 AGGTGGATGTTTCCATTTTGGGTTTTAAATTTGATATATCCCTGATAGCTATAATTTGGGTGAT 3958

Db |||||||

170 TGATG-----TTTCATTTTGGTTTTAAATTTGTATATCCCTGATAGCTATAATTTGGGTGAT 116

QY 3959 AGAAATCTTTATATACATTTAGATGCAAGTCTCTTGYCGGATATACGTATTGAGATATTA 4018

Db |||||||

115 AGAAATCTTTATATACATTTAGATGCAAGTCTCTTGYCGGATATATGTTTATGAGATATTA 56

QY 4019 CACCTAGTCTGTGGCTTGACTGTTTCTTTATGCTTTTGTGATGAATAGAAGTTTT 4073

Db |||||||

55 CACCTAGTCTGTGGCTTGACTGTTTCTTTATGCTTTTGTGATGAATAGAAGTTTT 1

## RESULT 3

US-09-833-263-343/c

; Sequence 343, Application US/09833263

; Patent No. US20020110547A1

## GENERAL INFORMATION:

; APPLICANT: Wang, Aijun

; APPLICANT: Clapper, Jonathan D.

; APPLICANT: Stolk, John A.

; APPLICANT: Meagher, Madeleine J.

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND

; FILE REFERENCE: 210121.471C12

; CURRENT APPLICATION NUMBER: US/09/833,263

; CURRENT FILING DATE: 2001-04-10

; NUMBER OF SEQ ID NOS: 1093

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 343

; LENGTH: 646

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-833-263-343

Query Match	12.5%	Score 552.2;	DB 9;	Length 646;
Best Local Similarity	93.0%;	Pred. No. 5.8e-111;		
Matches 609;	Conservative 4;	Mismatches 32;	Indels 10;	Gaps 3;
Qy	3419	AAAAAGCTTCTCAGCGCTTCCCTAGGGAACAGAAATTTGGGTGAGCCCAATCTGCAATTTCTTACT	3478	
Db	645	AAAGCTTCTCAGCGCTTCCCTAGGGAACAGAAATTTGGGAGGCCCAATCTGCAATTTCTTACT	586	
Qy	3479	ACAGGCATTGAGACAGCTTAGATTATTTCGAAATATTATATAGAGAGTTATTGAACACTTAATTT	3538	
Db	585	ACAGGCATTGAGACAGCTTAGATTATTTCGAAATATTATATAGAGAGTTATTGAACACTTAATTT	526	
Qy	3539	ATGATAGTGGTATGACATTTGGATAGAACATGGGATATCTTTAGAAGTAGAATTTGACAGGGC	3598	
Db	525	ATGATAGTGGTATGACATTTGGATAGAACATGGGATATCTTTAGAAGTAGAATTTGACAGGGC	466	
Qy	3599	ATATTAGTTGATGAAATGGAGTCATTTTGAGTCTVTTTAATAGCCCATGATCATATAATTACCA	3658	
Db	465	ATATTAGTTGATGAAATGGAGTCATTTTGAGTCTCTTAATAGCCCATGATCATATAATTACCA	406	
Qy	3659	AGTCAAGCTGCTGGAACATATGTCCTCCATTTTACAGTTAAGGAATATATATGACACATT	3718	
Db	405	AGTCAAGCTGCTGGAACATATGTCCTCCATTTTACAGTTAAGGAATATATATGACACATT	346	
Qy	3719	AATATTGTTTCTGTCATGCCCCACAATCCCTTCTTAAGGAAGACTGCGCTACTATAGCAG	3778	
Db	345	AATATTGTTTCTGTCATGCCCCACAATCCCTTCTTAAGGAAGACTGCGCTACTATAGCAG	286	
Qy	3779	TTTTTATATTGTCACAAATTTATGAATATATGAAGAGAGTTCTGGTACCTCCTGTCCTTT	3838	
Db	285	TTTTTATATTGTCACAAATTTATGAATATATGAATATGA- GAGTTCCTGGTACCTCCTGTCCTTT	227	
Qy	3839	ACAAATATTGGGTGTGTCACGATTTTTCCTTTTAAACMTTCCCAATTCGGGTGTGT	3898	
Db	226	ACAAATATTGG- --- TGTGTGCAGTATTTTTCCTTTTAAACCATTCCAATCGGTGTGTAG	171	
Qy	3899	AGGTGGATGTTTCCATTTGGGTTTTAAATTTGTATATATCCCTGATAGCTATATAATTTGGGTCAT	3958	
Db	170	TGATG- ----TTTCAATTTGGTTTAAATTTGTATATCCCTGATAGCTATATAATTTGGGTCAT	116	
Qy	3959	AGAAATCTTTATACATTTCTAGATGCAAGTCTCTTGTGCGGATATACGTATTGAGATATTA	4018	
Db	115	AGAAATCTTTATACATTTCTAGATGCAATCTCTTGTGCGGATATATGTATTGAGATATTA	56	
Qy	4019	CACCTAGCTGTGCGCTTGACGTGTTTTCTTTATGCTTTTTTGATGAATAGAAAGTTTT	4073	
Db	55	CACCTAGCTGTGCGCTTGACGTGTTTTCTTTATGCTTTTTTGATGAATAGAAAGTTTT	1	

```

RESULT 4
US - 025-380-343/c
; Sequence 343, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skeiky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick Thomas S.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOUNDS FOR IM
; TITLE OF INVENTION: OF COLON CANCER
; FILE REFERENCE: 210121.471C14

```

```

, CURRENT APPLICATION NUMBER: US/10/025.380
, CURRENT FILING DATE: 2001-12-19
, NUMBER OF SEQ ID NOS: 1129
, SOFTWARE: FastSEQ for Windows Version 4.0
, SEQ ID NO 343
, LENGTH: 646
, TYPE: DNA
, ORGANISM: Homo sapiens
US-10-025-380-343

```

Query Match	12.5%	Score 552.2;	DB 13;	Length 646;
Best Local Similarity	93.0%;	Pred. No. 5.8e-111;		
Matches 609;	Conservative 4;	Mismatches 32;	Indels 10;	Gaps 3;
Qy	3419	AAAAGCTTCTCAGCCTTCCTAGGGAACAGAAAATTGGGTGAGCCAAATCTGCAAAATTTCTTACT	3478	
Db	645	AAAGCTTCTCAGCCTTCCTAGGGAACAGAAAATTGGGAGCCCAATCTGCAAAATTTCTTACT	586	
Qy	3479	ACAGGCATTGAGACCACTAGATTATTTGAAATATTTATAGAGAGTTATGAACACCTAAATTT	3538	
Db	585	ACAGGCATTGAGACCACTAGATTATTTGAAATATTTATAGAGAGTTATGAACACCTAAATTT	526	
Qy	3539	ATGATAGTGTATGACATTCGATAGAACATGGGATACTTTTAGAAGTAGAAATGACAGGCG	3598	
Db	525	ATGATAGTGTATGACATTCGATAGAACATGGGATACTTTTAGAAGTAGAAATGACAGGCG	466	
Qy	3599	ATATTAGTTGATGAAATGGAGTCATTTTGAGTCTTTTAAATAGCCATGTCATATAATTACCA	3658	
Db	465	ATATTAGTTGATGAAATGGAGTCATTTTGAGTCTTTTAAATAGCCATGTCATATAATTACCA	406	
Qy	3659	AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTTCACAGTTTAAAGGAATAATAATGACAGATT	3718	
Db	405	AGTGAAGCTGGTGGAAACATATGGTCTCCATTTTTCACAGTTTAAAGGAATAATAATGACAGATT	346	
Qy	3719	AATATTGTTTCTGTGTCATGCCCCCAATCCCTTTCTAAGGAAGAGTCGCCCTACTATAGCAG	3778	
Db	345	AATATTGTTTCTGTGTCATGCCCCCAATCCCTTTCTAAGGAAGAGTCGCCCTACTATAGCAG	286	
Qy	3779	TTTTTTATATTTGTCAATTTTATGAATATAATGAATAGGAGTTCTGTGACCTCTGTCTTT	3838	
Db	285	TTTTTTATATTTGTCAATTTTATGAATATAATGAATAGA - GAGTTCTGTGTACCTCTGTCTTT	227	
Qy	3839	ACAAATATTTGGGTGTGTCCAGTATTTTTCCTTTTAAACMMTCCCAATTCGGGTGTGT	3898	
Db	226	ACAAATATTTGG - - - - - TGTGTGCAGTATTTTTCCTTTTTTAAACCAATTCGAATCGGTGTGTAG	171	
Qy	3899	AGGTGGATGTTTCCATTTGGGTTTTAAATTTTGTATATATCCCTGTAGTCATATAAATTTGGGTTCAT	3958	
Db	170	TGATG - - - - - TTTTCATTTTGGTTTTTAAATTTTGTATATATCCCTGTAGTCATATAAATTTGGGTTCAT	116	
Qy	3959	AGAAATCTTTATACATTTCTAGATGCAAGTCTCTTGTGCGGATATACGTATGTAGATATTTA	4018	
Db	115	AGAAATCTTTATACATTTCTAGATGCAAAATCTCTTGTGCGGATATATGTATGTAGATATTTA	56	
Qy	4019	CACCTAGTCTGTGGCTTGACTGTTTTTCTTTATGTCCTTTTGTGATGAATAGAAAGTTTT	4073	
Db	55	CACCTAGTCTGTGGCTTGACTGTTTTTCTTTATGTCCTTTTGTGATGAATAGAAAGTTTT	1	

```

RESULT 5
US-09-969-034-1463/C
; Sequence 1463, Application US/09969034
; Publication No. US2004011068A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Ratle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poorima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thisgalingam, Arunthathi
; APPLICANT: Lewtas, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially

```

```
; TITLE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1463
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 546
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-1463

Query Match      12.4%; Score 544.8; DB 11; Length 561;
Best Local Similarity 99.3%; Pred. No. 2.3e-109;
Matches 557; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 483 CAGAAGATGCATCA-CCTGTGTCATGATCAGTGAAGTTACCAATTAAAAACAATATCT 541
Db      |||||||
QY 561 CAGAAGATGCATCNCCTTGTGTCATGATCAGTGAAGTTACCAATTAAACCAATATCT 502
Db      |||||||
QY 542 GCATCATTTTGGAGCTCAATTTGTAAGCCGCAATATTTTACTGAAATTCCTGAAGCA 601
Db      |||||||
QY 501 GCATCATTTTGGAGCTCAATTTGTAAGCCGCAATATTTTACTGAAATTCCTGAAGCA 442
Db      |||||||
QY 602 GTTCAGTCCAAAGAGCAGCTCCACAAATTTGAAAGTTTTCACCACTCTTGATGAACCA 661
Db      |||||||
QY 441 GTTGAGTCCAAAGAGCAGCTCCACAAATTTGAAAGTTTTCACCACTCTTGATGAACCA 382
Db      |||||||
QY 662 TCTATTGGAAGTAAAAATTTGATCTGTGTCAGGACGCGAGGAAGAAACAAATCTTCAAA 721
Db      |||||||
QY 381 TCTATTGGAAGTAAAAATTTGATCTGTGTCAGGACGCGAGGAAGAAACAAATCTTCAAA 322
Db      |||||||
QY 722 GGGAAACAATTTATATTTTGAATGCGCAACAGCATAAGAAATTTGATTTCCGAGTTGTC 781
Db      |||||||
QY 321 GGGAAACAATTTATATTTTGAATGCGCAACAGCATAAGAAATTTGATTTCCGAGTTGTC 262
Db      |||||||
QY 782 TTTGGAGTGGGGAAGCTAGTTGATAACAGAGAGAAATGAAGAAACAATAATTTCTTT 841
Db      |||||||
QY 261 TTTGGAGTGGGGAAGCTAGTTGATAACAGAGAGAAATGAAGAAACAATAATTTCTTT 202
Db      |||||||
QY 842 TTGGCTCCGGGAACGTGTGTGTGTGATACAGGAATAACAACTCACAGACCTTAAATTCCT 901
Db      |||||||
QY 201 TTGGCTCCGGGAACGTGTGTGTGTGATACAGGAATAACAACTCACAGACCTTAAATTCCT 142
Db      |||||||
QY 902 GACTGTCAAGAAATGGATTCAGTCAATAATGGATATGCTCCAAAGGCAAGTCTTAGA 961
Db      |||||||
QY 141 GACTGTCAAGAAATGGATTCAGTCAATAATGGATATGCTCCAAAGGCAAGTCTTAGA 82
Db      |||||||
QY 962 CCTATTCTTGAAGCAAAATTTGATTTGGGCTGATTTTCATGACTACAAAGAAATTAATCT 1021
Db      |||||||
QY 81 CCTATTCTTGAAGCAAAATTTGATTTGGGCTGATTTTCATGACTACAAAGAAATTAATCT 22
Db      |||||||
QY 1022 GATCCTCAGGGCCATCCCACT 1042
Db      |||||||
QY 21 GATCCTCAGGGCCATCCCACT 1
Db      |||||||

RESULT 6
US-09-878-178-1737
; Sequence 1737, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0

; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1737
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-878-178-1737

Query Match      12.2%; Score 537.4; DB 9; Length 541;
Best Local Similarity 99.4%; Pred. No. 9.5e-108;
Matches 538; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3130 ACCTGTGATCCAGCAAGAGGAGTTCCAGTCAAGAGTCACCTACAACTGATTTAGTTT 3189
Db      |||||||
QY 3190 AGAGAAATGAGAAATCGAACAGTGAAGAAATGAGAGCCATATTTCCATGACTTCCCTTTGTA 3249
Db      |||||||
QY 61 AGAGAAATGAGAAATCGAACAGTGAAGAAATGAGAGCCATATTTCCATGACTTCCCTTTGTA 120
Db      |||||||
QY 3250 ACAGAAAGCAACAGAAAGGAGCAAGAGGCTGGCCCTCTACATCACTCTCACCTTCCAAATCTT 3309
Db      |||||||
QY 121 ACAGAAAGCAACAGAAAGGAGCAAGAGGCTGGCCCTCTACATCACTCTCACCTTCCAAATCTT 180
Db      |||||||
QY 3310 GTGGAAGTGCATCTACTTGCAGAACCAAAATTAACATTACTTCCAAAGTTCTGGCTGCTTGC 3369
Db      |||||||
QY 181 GTGGAAGTGCATCTACTTGCAGAACCAAAATTAACATTACTTCCAAAGTTCTGGCTGCTTGC 240
Db      |||||||
QY 3370 AGGTGGAACTCCAGCTGCAAGGGAGTTAGGGAATGAAGGCTCTTTTAAAGCTTCTC 3429
Db      |||||||
QY 241 AGGTGGAACTCCAGCTGCAAGGGAGTTAGGGAATGAAGGCTCTTTTAAAGCTTCTC 300
Db      |||||||
QY 3430 AGCCTTCTTAGGGAACAGAAATTTGGGTGAGCCAAATCTGCAATTTCTACTACAGGCAATGA 3489
Db      |||||||
QY 301 AGCCTTCTTAGGGAACAGCAATTTGGGTGAGCCAAATCTGCAATTTCTACTACAGGCAATGA 360
Db      |||||||
QY 3490 GACCAAGTATGATTTGAAATATTAATAGAGTTATGAACACTTAAATATGATAGTGT 3549
Db      |||||||
QY 361 GACCAAGTATGATTTGAAATATTAATAGAGTTATGAACACTTAAATATGATAGTGT 420
Db      |||||||
QY 3550 ATGACATTTGGATAGAACATGGGATCTTTAGAGTAGAATTCAGAGGCAATATAGTTGA 3609
Db      |||||||
QY 421 ATGACATTTGGATAGAACATGGGATCTTTAGAGTAGAATTCAGAGGCAATATAGTTGA 480
Db      |||||||
QY 3610 TGAAATGGAGTCATTTGAGTCTTTTAAATAGCCATGTATCATAAATTTACCAAGTGAAGCTGG 3669
Db      |||||||
QY 481 TGAAATGGAGTCATTTGAGTCTTTTAAATAGCCATGTATCATAAATTTACCAAGTGAAGCTGG 540
Db      |||||||
QY 3670 T 3670
Db      |||||||
QY 541 T 541
Db      |||||||

RESULT 7
US-10-046-935-1737
; Sequence 1737, Application US/10046935
; Publication No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 1737
; LENGTH: 541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-046-935-1737

```

Query Match 12.2%; Score 537.4; DB 13; Length 541;  
Best Local Similarity 99.4%; Pred. No. 9.5e-108;  
Matches 538: Conservative 1; Mismatches 2; Indels 0;

Qy	3130	ACCTGTGATCCAGCAAGAAAGGGAGTTC	CAGTCAAGAGTCACTCAACTGATTAGTTGTTT	3189
Dd	1	ATCTGTGATCCAGCAAGAAAGGGAGTTC <td>CAGTCAAGAGTCACTCAACTGATTAGTTGTTT</td> <td>60</td>	CAGTCAAGAGTCACTCAACTGATTAGTTGTTT	60

Qy	3190	AGAGAAATGAGAAATGGAAACAGTGGAGGAATGGAGGCCATATTTCCATGACTTCCCTTGTA	3249
Db	61	AGAGAAATGAGAAATGGAAACAGTGGAGGAATGGAGGCCATATTTCCATGACTTCCCTTGTA	120

Qy	3250	ACAGAAGCAACAGAAAGGACAAAGAGGTGGCGCTTACATCACTCTCACCTTCCAAATCTT	3309
Db	121	ACAGAAGCAACAGAAAGGACAAAGAGGTGGCGCTTACATCACTCTCACCTTCCAAATCTT	180

Qy 3310 GTGGAAGTGCATCTACTTGTCCAGAACCAAAATAACTTATCTCCAGTTCTGGTGTCTTGC 3369

Dh 181 GTGGGAGTGCATCTACTTGTCCAGAACCAAAATAACTTATCTCCAGTTCTGGTGTCTTGC 240

QY 3370 AGGTGGAACTCCAGCTGCAAGGAGTGTAGGGAAATGAAGGCTCTTTTAAAGCTTCTC 3429

[illegible]

3490 GACCAGTTAGATTATTGAAATATTATAGAGAGTTATGAACACTTAAATTTATGATAGTCGT 3549

3550 ATGACATTGGATAGAACATGGGATCTTTAGAGTAGAATTGCAGGCCATATTAGTTGA 3609

Qy 3610 TGAATGGAGTCATTGAGTCCTTTAATTAGCCATGTATCATTAATACCAAGTGAAGCTGG 3669

Qy 3670 T 3670 |

## RESULT 8

; Sequence 1737, Application US/10146502  
 ; Publication No. US20030069180A1  
 ; GENERAL INFORMATION:

: Wang, Liqun  
: APPLICANT: Harlocker, Susan L.  
: APPLICANT: Secrist, Heather  
: APPLICANT: Wang, Aijun

REF: 1  
; 2  
; 3  
; 4  
; 5  
; 6  
; 7  
; 8  
; 9  
; 10  
; 11  
; 12  
; 13  
; 14  
; 15  
; 16  
; 17  
; 18  
; 19  
; 20  
; 21  
; 22  
; 23  
; 24  
; 25  
; 26  
; 27  
; 28  
; 29  
; 30  
; 31  
; 32  
; 33  
; 34  
; 35  
; 36  
; 37  
; 38  
; 39  
; 40  
; 41  
; 42  
; 43  
; 44  
; 45  
; 46  
; 47  
; 48  
; 49  
; 50  
; 51  
; 52  
; 53  
; 54  
; 55  
; 56  
; 57  
; 58  
; 59  
; 60  
; 61  
; 62  
; 63  
; 64  
; 65  
; 66  
; 67  
; 68  
; 69  
; 70  
; 71  
; 72  
; 73  
; 74  
; 75  
; 76  
; 77  
; 78  
; 79  
; 80  
; 81  
; 82  
; 83  
; 84  
; 85  
; 86  
; 87  
; 88  
; 89  
; 90  
; 91  
; 92  
; 93  
; 94  
; 95  
; 96  
; 97  
; 98  
; 99  
; 100  
; 101  
; 102  
; 103  
; 104  
; 105  
; 106  
; 107  
; 108  
; 109  
; 110  
; 111  
; 112  
; 113  
; 114  
; 115  
; 116  
; 117  
; 118  
; 119  
; 120  
; 121  
; 122  
; 123  
; 124  
; 125  
; 126  
; 127  
; 128  
; 129  
; 130  
; 131  
; 132  
; 133  
; 134  
; 135  
; 136  
; 137  
; 138  
; 139  
; 140  
; 141  
; 142  
; 143  
; 144  
; 145  
; 146  
; 147  
; 148  
; 149  
; 150  
; 151  
; 152  
; 153  
; 154  
; 155  
; 156  
; 157  
; 158  
; 159  
; 160  
; 161  
; 162  
; 163  
; 164  
; 165  
; 166  
; 167  
; 168  
; 169  
; 170  
; 171  
; 172  
; 173  
; 174  
; 175  
; 176  
; 177  
; 178  
; 179  
; 180  
; 181  
; 182  
; 183  
; 184  
; 185  
; 186  
; 187  
; 188  
; 189  
; 190  
; 191  
; 192  
; 193  
; 194  
; 195  
; 196  
; 197  
; 198  
; 199  
; 200  
; 201  
; 202  
; 203  
; 204  
; 205  
; 206  
; 207  
; 208  
; 209  
; 210  
; 211  
; 212  
; 213  
; 214  
; 215  
; 216  
; 217  
; 218  
; 219  
; 220  
; 221  
; 222  
; 223  
; 224  
; 225  
; 226  
; 227  
; 228  
; 229  
; 230  
; 231  
; 232  
; 233  
; 234  
; 235  
; 236  
; 237  
; 238  
; 239  
; 240  
; 241  
; 242  
; 243  
; 244  
; 245  
; 246  
; 247  
; 248  
; 249  
; 250  
; 251  
; 252  
; 253  
; 254  
; 255  
; 256  
; 257  
; 258  
; 259  
; 260  
; 261  
; 262  
; 263  
; 264  
; 265  
; 266  
; 267  
; 268  
; 269  
; 270  
; 271  
; 272  
; 273  
; 274  
; 275  
; 276  
; 277  
; 278  
; 279  
; 280  
; 281  
; 282  
; 283  
; 284  
; 285  
; 286  
; 287  
; 288  
; 289  
; 290  
; 291  
; 292  
; 293  
; 294  
; 295  
; 296  
; 297  
; 298  
; 299  
; 300  
; 301  
; 302  
; 303  
; 304  
; 305  
; 306  
; 307  
; 308  
; 309  
; 310  
; 311  
; 312  
; 313  
; 314  
; 315  
; 316  
; 317  
; 318  
; 319  
; 320  
; 321  
; 322  
; 323  
; 324  
; 325  
; 326  
; 327  
; 328  
; 329  
; 330  
; 331  
; 332  
; 333  
; 334  
; 335  
; 336  
; 337  
; 338  
; 339  
; 340  
; 341  
; 342  
; 343  
; 344  
; 345  
; 346  
; 347  
; 348  
; 349  
; 350  
; 351  
; 352  
; 353  
; 354  
; 355  
; 356  
; 357  
; 358  
; 359  
; 360  
; 361  
; 362  
; 363  
; 364  
; 365  
; 366  
; 367  
; 368  
; 369  
; 370  
; 371  
; 372  
; 373  
; 374  
; 375  
; 376  
; 377  
; 378  
; 379  
; 380  
; 381  
; 382  
; 383  
; 384  
; 385  
; 386  
; 387  
; 388  
; 389  
; 390  
; 391  
; 392  
; 393  
; 394  
; 395  
; 396  
; 397  
; 398  
; 399  
; 400  
; 401  
; 402  
; 403  
; 404  
; 405  
; 406  
; 407  
; 408  
; 409  
; 410  
; 411  
; 412  
; 413  
; 414  
; 415  
; 416  
; 417  
; 418  
; 419  
; 420  
; 421  
; 422  
; 423  
; 424  
; 425  
; 426  
; 427  
; 428  
; 429  
; 430  
; 431  
; 432  
; 433  
; 434  
; 435  
; 436  
; 437  
; 438  
; 439  
; 440  
; 441  
; 442  
; 443  
; 444  
; 445  
; 446  
; 447  
; 448  
; 449  
; 450  
; 451  
; 452  
; 453  
; 454  
; 455  
; 456  
; 457  
; 458  
; 459  
; 460  
; 461  
; 462  
; 463  
; 464  
; 465  
; 466  
; 467  
; 468  
; 469  
; 470  
; 471  
; 472  
; 473  
; 474  
; 475  
; 476  
; 477  
; 478  
; 479  
; 480  
; 481  
; 482  
; 483  
; 484  
; 485  
; 486  
; 487  
; 488  
; 489  
; 490  
; 491  
; 492  
; 493  
; 494  
; 495  
; 496  
; 497  
; 498  
; 499  
; 500  
; 501  
; 502  
; 503  
; 504  
; 505  
; 506  
; 507  
; 508  
; 509  
; 510  
; 511  
; 512  
; 513  
; 514  
; 515  
; 516  
; 517  
; 518  
; 519  
; 520  
; 521  
; 522  
; 523  
; 524  
; 525  
; 526  
; 527  
; 528  
; 529  
; 530  
; 531  
; 532  
; 533  
; 534  
; 535  
; 536  
; 537  
; 538  
; 539  
; 540  
; 541  
; 542  
; 543  
; 544  
; 545  
; 546  
; 547  
; 548  
; 549  
; 550  
; 551  
; 552  
; 553  
; 554  
; 555  
; 556  
; 557  
; 558  
; 559  
; 560  
; 561  
; 562  
; 563  
; 564  
; 565  
; 566  
; 567  
; 568  
; 569  
; 570  
; 571  
; 572  
; 573  
; 574  
; 575  
; 576  
; 577  
; 578  
; 579  
; 580  
; 581  
; 582  
; 583  
; 584  
; 585  
; 586  
; 587  
; 588  
; 589  
; 590  
; 591  
; 592  
; 593  
; 594  
; 595  
; 596  
; 597  
; 598  
; 599  
; 600  
;

```

;
; CURRENT APPLICATION NUMBER: 05/10/146,502
;
; CURRENT FILING DATE: 2002-05-14
;
; NUMBER OF SEQ ID NOS: 2241
;
; SOFTWARE: FastSeq for Windows Version 4.0
;

```

```

? SEQ ID NO 1737
? LENGTH: 541
? TYPE: DNA
? ORGANISM: Homo sapiens

```

US-10-146-502-1737

Query Match 12.2%; Score 537.4; DB 14; Length 541;  
Best Local Similarity 99.4%; Pred. No. 9.5e-108;  
Matches 538; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

3130 ACCCTGGATCCAGCAGGAGGAGTCCAGTCCAGTCAAGAGTCACTACAACCTGATTGTTT 60

5190 AGAGAAATGAGAAATGGAAACAGTGGAGAAATGGAGCCCATATTTCCATGACATCCCTTCTGTAA 5240  
61 AGAGAAATGAGAAATGGAAACAGTGGAGAAATGGAGCCCATATTTCCATGACATCCCTTCTGTAA 120

3250 ACAGAGGCAACAGAGAGGGACAGAGAGGCTGGCCCTCTACATCACCTCTCCACCTTCCAATCTT 180

3310 GTGGGAGTGCATCTACTTCCAGAACCAAAATTAATCTCAAGTTCCTGGCTGGCTTGC 3360  
181 GTGGAAGTGCATCTACTTCCAGAACCAAAATTAATCTCAAGTTCCTGGCTGGCTTGC 240

3370 AGGTGGAACTCCAGCTCAGGGAGTTAGGGAAATGAAGTCTTTTTTAAAGCTTCTC 300

3430 AGCCTTCTCTAGGGGACAGAAAATGGGTGAGCCCAATCTGCAATTTCTACTACAGGCATTTGA 3488  
301 AGCCTTCTCTAGGGGACAGACATTTGGGTGAGCCCAATCTGCAATTTCTACTACAGGCATTTGA 360

3490 GACCAAGTTAGATTATTGAAATATATTATAGAGAGTTATGAACACCTTAAATTATGATAGTGGT 3540  
361 GACCAAGTTAGATTATTGAAATATATTATAGAGAGTTATGAACACCTTAAATTATGATAGTGGT 420

3550 ATGACATTGGATAGAACATGGGATACCTTTAGAGTAGAATTGACAGGGCATATTAGTTGA 3609

421 ATGACATTGGATAGAACATGGGATACCTTTAGAGTAGAATTGACAGGGCATATTAGTTGA 480

3610 TGAATGGAGTCATTGTGAGTCCTTTAATAGCCATGTATACATAATTACCAAGTGAAGCTGG 3669

481 TGAATGGAGTCATTGTGAGTCCTTTAATAGCCATGTATACATAATTACCAAGTGAAGCTGG 540

3670 T 3670  
—  
541 T 541

RESULT 9  
S-10-102-524-532

Publication No. US20030109434A1  
GENERAL INFORMATION:  
APPLICANT: Algate, Paul A.

APPLICANT: Gaiger, Alexander  
APPLICANT: Gordon, Brian  
APPLICANT: Harlocker, Susan L.

FILE OF INVENTOR: COLLEAGUES AND FRIENDS OF THE  
TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER  
FILE REFERENCE: 210121.572  
CURRENT APPLICATION NUMBER: US/10/102.524

```

CURRENT FILING DATE: 2002-03-13
NUMBER OF SEQ ID NOS: 1863
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 532

```

LENGTH: 507  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 S-10-102-524-532

Query Match 11.5%; Score 507; DB 15; Length 587;  
Best Local Similarity 93.6%; Pred. No. 4.9e-101;  
Matches 550; Conservative 4; Mismatches 24; Indels 10; Gaps 3;

3419	Qy	AAAAGCTTCTCAGCCTTCCTAGGAAACAGAAATTTGGGTGAGCCAATCTGCAATTTCTTACT	3478
1	Db	AAAAGCTTCTCAGCCTTCCTAGGAAACAGAAATTTGGGTGAGCCAATCTGCAATTTCTTACT	60
3479	Qy	ACGGCAATCGAGACCAAGTAGATTATTTGAAATATTTATAGAGAGTTATGAACACTTAAATT	3538
61	Db	ACGGCAATCGAGACCAAGTTAGATTATTTGAAATATTTATAGAGAGTTATGAACACTTAAATT	120
3539	Qy	ATGATAGTGGTATGACATTTGGATAGAACATGGGATACTTTTAGAAGTAGAATTTGACAGGGC	3598
121	Db	ATGATAGTGGTATGACATTTGGATAGAACATGGGATACTTTTAGAAGTAGAATTTGACAGGGC	180
3599	Qy	ATATTAGTTGATGAAATGGAGTCATTTTGAGTCTTATTAATAGCCATGATCATTAATTACCA	3658
181	Db	ATATTAGTTGATGAAATGGAGTCATTTTGAGTCTTATTAATAGCCATGATCATTAATTACCA	240
3659	Qy	AGTCAGCTGGTGAAACATATGGTCTCCATTTTACAGTTAAGGAATATATGACACAGATT	3718
241	Db	AGTCAGCTGGTGAAACATATGGTCTCCATTTTACAGTTAAGGAATATATGAGACAGATT	300
3719	Qy	AATATTGTTTCTGTCATGCCACAAATCCCTTTCTAAGGAAGAAGCTGCCCTACTATAGCAG	3778
301	Db	AATATTGTTTCTGTCATGCCACAAATCCCTTTCTAAGGAAGAAGCTGCCCTACTATAGCAG	360
3779	Qy	TTTTTATATTGTCAAATTTATGAATATAATGAATGAGAGGTTCTGGTACCTCTGTCTTT	3838
361	Db	TTTTTATATTGTCAAATTTATGAATATAATGAATGA - GAGTTCTGGTACCTCTGTCTTT	419
3839	Qy	ACAAATATTGGGTGTTGCCAGTATTTTTCCCTTTTTTAACTCCCAATTCGGGTGCTGT	3898
420	Db	ACAAATATTGG - - - TGGTGTGAGTATTTTTCCTTTTTTAACTCCCAATTCGGGTGCTGTAG	475
3899	Qy	AGGTGGGATGTTTCCATTTGGGTGTTTTTAATTTGTATATCCCTGATAGCTATAATTGGGTGAT	3958
476	Db	TGATG - - - -TTTTCATTTTGGTTTTAATTTGTATATCCCTGATAGCTATAATTGGGTGAT	530
3959	Qy	AGAAATCTTTTATACATCTAGATGCAAGTCTCTTGYCGGNATATACGTATTGAGATA	4015
531	Db	AGAAATCTTTTATACATCTAGATGCAAGTCTCTTGYCGGNATATATGATTGAGATA	587

```

RESULT 10
US-10-242-5355A-26197
; Sequence 26197, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: Chondrogene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26197
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-5355A-26197

```

Query Match 10.9%; Score 478.4; DB 17; Length 481;

	Best Local Similarity	99.6%; Matches	99.6%; Conservative	Pred. No. 8.7e-95; 0; Mismatches	0; Indels	0; Gaps	0;
QY	3126	ATGTACCTGTGATCCAGCAAGAAGGAGT	TCCAGTCAAGAGTCACTACA	CTGATTAGTT	3185		
Db	1	ATGTACCTGTGATCCAGCAAGT	TTCAGTCAAGAGTCACTACA	CTGATTAGTT	60		
QY	3186	GTTTATGAGATGAGAAATGGAA	CAGTGGAGTGGAGGCCATATTTCCAT	GTACTTCGCTT	3245		
Db	61	GTTTATGAGATGAGAAATGGAA	CAGTGGAGTGGAGGCCATATTTCCAT	GTACTTCGCTT	120		
QY	3246	GTAACAGAAAGCAACAGAGGGCA	AGAGGCTGGCCCTCATCATCTCACCTT	CCAAA	3305		
Db	121	GTAACAGAAAGCAACAGAGGGCA	AGAGGCTGGCCCTCATCATCTCACCTT	CCAAA	180		
QY	3306	TCTTGTGGAAGTGCACTCTCTG	CCAGAACCAAAATTAACCTTAC	TCGAGTTC	3365		
Db	181	TCTTGTGGAAGTGCACTCTCTG	CCAGAACCAAAATTAACCTTAC	TCGAGTTC	240		
QY	3366	TTGCGAGTGGAACTCCAGCTGCA	AGGAGTTAGGGAATCAAGGTCT	TTTTTAAAGCT	3425		
Db	241	TTGCGAGTGGAACTCCAGCTGCA	AGGAGTTAGGGAATCAAGGTCT	TTTTTAAAGCT	300		
QY	3426	TCTCAGGCTTCCTAGGGAAACAGA	AAATTTGGGTGAGCCAAATCTG	CAATTTCTACTACAGGCA	3485		
Db	301	TCTCAGGCTTCCTAGGGAAACAGA	AAATTTGGGTGAGCCAAATCTG	CAATTTCTACTACAGGCA	360		
QY	3486	TTGAGACCAAGTTAGATTATTTG	AAATATTTATAGAGAGTTATGA	CACCTTAATTTATGATAG	3545		
Db	361	TTGAGACCAAGTTAGATTATTTG	AAATATTTATAGAGAGTTATGA	CACCTTAATTTATGATAG	420		
QY	3546	TGGTATCACATTTGGATGAACAT	GGGATACCTTTTAGAGTAGAAT	TGACAGGGCATATTAG	3605		
Db	421	TGGTATCACATTTGGATGAACAT	GGGATACCTTTTAGAGTAGAAT	TGACAGGGCATATTAG	480		
QY	3606	T	3606	T	3606		
Db	481	T	481	T	481		

```

RESULT 11
US-10-085-783A-26197
; Sequence 26197, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26197
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-26197

```

Query Match 10.9%; Score 478.4; DB 18; Length 481;  
Best Local Similarity 99.6%; Pred. No. 8.7e-95;  
Matches 479; Conservative 0; Mismatches 2; Indels 0;

QY 3126 ATGTACTGTGATCCAGCAAGAGGAGTTCCAGTCAAGAGTCACTACAACTGATTAGTT 3185  
Db 1 ATGTACTGTGATCCAGCAAGAGGAGTTCCAGTCAAGAGTCACTACAACTGATTAGTT 60  
QY 3186 GTTTAGAGATGAGAATGCAAGTCAAGAGTGGAGGCCATATTTCCATGACTTCCCTT 3245  
Db 61 GTTTAGAGATGAGAATGCAAGTCAAGAGTGGAGGCCATATTTCCATGACTTCCCTT 120  
QY 3246 GTAACAGAAAGCAACAGAGGCAAGAGGCTGGCCCTCTACATCACTCTCACCTTCCAAA 3305  
Db 121 GTAACAGAAAGCAACAGAGGCAAGAGGCTGGCCCTCTACATCACTCTCACCTTCCAAA 180  
QY 3306 TCTTGTGGAAGTGCATCTACTTCCAGAACCAATTAACCTTCTTCCAGTCTGGCTG 3365  
Db 181 TCTTGTGGAAGTGCATCTACTTCCAGAACCAATTAACCTTCTTCCAGTCTGGCTG 240  
QY 3366 TTGAGGTGGAATCCAGCTGCAAGGAGTTAGGAAATCAAGTCTTTTTTAAAGCT 3425  
Db 241 TTGAGGTGGAATCCAGCTGCAAGGAGTTAGGAAATCAAGTCTTTTTTAAAGCT 300  
QY 3426 TCTCAGCCTTCTTAGGAAACAGAAATTTGGGTGAGCCCAATCTGCAATTTCTACTACAGGCA 3485  
Db 301 TCTCAGCCTTCTTAGGAAACAGAAATTTGGGTGAGCCCAATCTGCAATTTCTACTACAGGCA 360  
QY 3486 TTGAGACCAGTTAGATTTTGAATATTTATAGAGAGTTATGAACACTTAAATTTATGATAG 3545  
Db 361 TTGAGACCAGTTAGATTTTGAATATTTATAGAGAGTTATGAACACTTAAATTTATGATAG 420  
QY 3546 TGTATGACATTCGATAGACATGGGATACCTTTAGAGTAGAATTCAGAGGCGATATTAG 3605  
Db 421 TGTATGACATTCGATAGACATGGGATACCTTTAGAGTAGAATTCAGAGGCGATATTAG 480  
QY 3606 T 3606  
Db 481 T 481

RESULT 12  
US-10-242-535A-7179  
; Sequence 7179, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 7179  
; LENGTH: 439  
; TYPE: DNA  
; ORGANISM: Human  
US-10-242-535A-7179  
  
Query Match 10.0%; Score 438.2; DB 17; Length 439;  
Best Local Similarity 99.3%; Pred. No. 5.9e-86;  
Matches 437; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3344 CTTACTTCCAAGTTCTGGCTGCTTGCAGGTGGAACTCCAGCTCAAGGGAGTTAGGGAAA 3403  
Db 1 CTTACTTCCAAGTTCTGGCTGCTTGCAGGTGGAACTCCAGCTCAAGGGAGTTAGGGAAA 60  
QY 3404 TGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTCAGGCAACAGAAATTTGGGTGAGCCAA 3463  
Db 61 TGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTCAGGCAACAGAAATTTGGGTGAGCCAA 120  
QY 3464 TCTGCAATTTCTACTACAGGCATTTGAGACCATTTAGATTTATGAAATATTTATAGAGTT 3523  
Db 121 TCTGCAATTTCTACTACAGGCATTTGAGACCATTTAGATTTATGAAATATTTATAGAGTT 180  
QY 3524 ATGAACACTTAAATTTATGATAGTGGTATGCATTTGGATAGAACATCGGATACCTTTAGAAG 3583  
Db 181 ATGAACACTTAAATTTATGATAGTGGTATGCATTTGGATAGAACATCGGATACCTTTAGAAG 240

Db 61 TGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTCAGGCAACAGAAATTTGGGTGAGCCAA 120  
QY 3464 TCTGCAATTTCTACTACAGGCATTTGAGACCATTTAGATTTATGAAATATTTATAGAGTT 3523  
Db 121 TCTGCAATTTCTACTACAGGCATTTGAGACCATTTAGATTTATGAAATATTTATAGAGTT 180  
QY 3524 ATGAACACTTAAATTTATGATAGTGGTATGCATTTGGATAGAACATCGGATACCTTTAGAAG 3583  
Db 181 ATGAACACTTAAATTTATGATAGTGGTATGCATTTGGATAGAACATCGGATACCTTTAGAAG 240  
QY 3584 TAGAATTCAGGGCATATTTAGTGAATGAAATGGAGTCAATTTGAGTCTTCTTAATAGCCAT 3643  
Db 241 TAGAATTCAGGGCATATTTAGTGAATGAAATGGAGTCAATTTGAGTCTTCTTAATAGCCAT 300  
QY 3644 GTATCATATTAACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAA 3703  
Db 301 GTATCATATTAACCAAGTGAAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAA 360  
QY 3704 TATAATGGACAGATTAATTTGTTTCTGTCTATGCCCAATCCCTTTCTAAGGAAGACT 3763  
Db 361 TATAATGGACAGATTAATTTGTTTCTGTCTATGCCCAATCCCTTTCTAAGGAAGACT 420  
QY 3764 GCCTACTATAGCAGTTT 3782  
Db 421 GCCTACTATAGCAGTTT 439

RESULT 13  
US-10-085-783A-7179  
; Sequence 7179, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 7179  
; LENGTH: 439  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-7179

Query Match 10.0%; Score 438.2; DB 18; Length 439;  
Best Local Similarity 99.5%; Pred. No. 5.9e-86;  
Matches 437; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3344 CTTACTTCCAAGTTCTGGCTGCTTGCAGGTGGAACTCCAGCTCAAGGGAGTTAGGGAAA 3403  
Db 1 CTTACTTCCAAGTTCTGGCTGCTTGCAGGTGGAACTCCAGCTCAAGGGAGTTAGGGAAA 60  
QY 3404 TGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTCAGGCAACAGAAATTTGGGTGAGCCAA 3463  
Db 61 TGAAGGTCTTTTTTAAAGCTTCTCAGCCTTCTCAGGCAACAGAAATTTGGGTGAGCCAA 120  
QY 3464 TCTGCAATTTCTACTACAGGCATTTGAGACCATTTAGATTTATGAAATATTTATAGAGTT 3523  
Db 121 TCTGCAATTTCTACTACAGGCATTTGAGACCATTTAGATTTATGAAATATTTATAGAGTT 180  
QY 3524 ATGAACACTTAAATTTATGATAGTGGTATGCATTTGGATAGAACATCGGATACCTTTAGAAG 3583  
Db 181 ATGAACACTTAAATTTATGATAGTGGTATGCATTTGGATAGAACATCGGATACCTTTAGAAG 240

QY 3584 TAGAATTGACGGCATATTAGTTGATGAATGGAGTCAATTTGAGTCTVTTAATAGCCAT 3643  
Db |||||  
QY 241 TAGAATTGACGGCATATTAGTTGATGAATGGAGTCAATTTGAGTCTCTTAATAGCCAT 300  
Db |||||  
QY 3644 GTATCATTAATTACCAAGTGAAGCTGCTGGAACATATGGTCTCCATTTTACAGTTAAGGAA 3703  
Db |||||  
QY 301 GTATCATTAATTACCAAGTGAAGCTGCTGGAACATATGGTCTCCATTTTACAGTTAAGGAA 360  
Db |||||  
QY 3704 TATAATGGACAGATTAATATTGTTTCTGTATGCCCAATGCCCTTCTAAGGAAGACT 3763  
Db |||||  
QY 361 TATAATGGACAGATTAATATTGTTTCTGTATGCCCAATGCCCTTCTAAGGAAGACT 420  
Db |||||  
QY 3764 GCCCTACTATAGCAGTTT 3782  
Db |||||  
QY 421 GCCCTACTATAGCAGTTT 439  
Db |||||  
RESULT 14  
US-09-918-995-13834  
; Sequence 13834, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FROM VARIOUS CDNA LIBRARIES  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 13834  
; LENGTH: 468  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(468)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-13834

Query Match 9.2%; Score 405.6; DB 10; Length 468;  
Best Local Similarity 97.4%; Pred. No. 9.2e-79;  
Matches 411; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 2938 TGTCTTACTGAGCTGGGTGGTCTGTTTGGAGCTTATTAGAGTCTCTAGTTTTCCTA 2997  
Db |||||  
QY 47 TCGTCTTACTGAGCTGGGTGGTCTGTTTGGAGCTTATTAGAGTCTCTAGTTTTCCTA 106  
Db |||||  
QY 2998 CTTTATAAGTAGAATGGTGAGATTGTTTCTTTTCTACCTAAAGGGAGATGGTAAGA 3057  
Db |||||  
QY 107 CTTTATAAGCAGAAACGGGAGATTGTTTGTGTTTCTACCTTAAAGGGAGATGGCAAGA 166  
Db |||||  
QY 3058 AACAATGAATGCTTTTTTCAAACTTATTGACAAGTGATTTTCAAGTCTGTGTTCAAAA 3117  
Db |||||  
QY 167 AACATGAATGCTTTTTTCAAACTTATTGACGAGTGATTTTCAAGTCTGTGTTCAAAA 226  
Db |||||  
QY 3118 ATATATTCTATGTACCTGTGATCCAGCAAGAGGAGTTCAGTCAAGAGTCACTACAAT 3177  
Db |||||  
QY 227 ATATATTCTATGTACCTGTGATCCAGCAAGAGGAGTTCAGTCAAGAGTCACTACAAT 286  
Db |||||  
QY 3178 GATTAGTTGTTTAGAATGAGAAATGGAACAGTGAAGATGGAGGCCATATTTCCATGA 3237  
Db |||||  
QY 287 GATTAGTTGTTTAAAGAAATGAGAAATGGAACAGTGAAGATGGAGGCCATATTTCCATGA 346  
Db |||||  
QY 3238 CTTCCCTTGTAAACAGAACCAAGAGGACAAAGAGGCTGGCTCTACATCACTCTCAC 3297  
Db |||||  
QY 347 CTTCCCTTGTAAACAGAACCAAGAGGACAAAGAGGCTGGCTCTACATCACTCTCAC 406  
Db |||||  
QY 3298 CTTCCAAATCTTGTGGAAGTGCACTTACTTGCAGAACCAAAATTAATCTTACTTCCAGTT 3357  
Db |||||  
QY 407 CTTCCAAATCTTGTGGAAGTGCACTTACTTGCAGAACCAAAATTAATCTTACTTCCAGTT 466  
Db |||||

QY 3358 CT 3359  
Db |||||  
QY 467 CT 468  
Db |||||  
RESULT 15  
US-09-969-034-1974/c  
; Sequence 1974, Application US/09969034  
; Publication No. US20040110668A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess, Christopher C.  
; APPLICANT: Astle, Jon H.  
; APPLICANT: Carroll, Eddie III  
; APPLICANT: Catino, Theodore J.  
; APPLICANT: Dwivedi, Poornima  
; APPLICANT: Molino, Gary A.  
; APPLICANT: Thiagalingam, Arunthathi  
; APPLICANT: Lewis, Marcia E.  
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially  
; Expressed in Cancer Tissue  
; FILE REFERENCE: 1657/1032  
; CURRENT APPLICATION NUMBER: US/09/969,034  
; CURRENT FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: 60/237,271  
; PRIOR FILING DATE: 2000-02-10  
; NUMBER OF SEQ ID NOS: 4494  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1974  
; LENGTH: 483  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 36, 41, 58, 70, 85, 118, 128, 133, 148, 152, 183, 187, 195,  
; LOCATION: 203, 211, 217, 247, 289, 301, 303, 315, 324, 331, 342, 346,  
; LOCATION: 358, 367, 378, 402, 405, 412, 417, 420, 427, 429, 432, 442,  
; LOCATION: 444, 470, 475, 477  
; OTHER INFORMATION: n = A,T,C or G  
US-09-969-034-1974

Query Match 8.2%; Score 362; DB 11; Length 483;  
Best Local Similarity 83.8%; Pred. No. 3.8e-69;  
Matches 383; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
QY 1253 CCCACTGTAAGAGTCTCTGCAAAACAAGCTCTAATAATAAGTAGTATGTTATCAATACT 1312  
Db |||||  
QY 473 CCCNCTGTAAGAGTCTCTGCAAAACAAGTNTAATAAANAGNANGGTATCNAANACT 414  
Db |||||  
QY 1313 TTGGCTTAAGATGAGATCCAAACTATCAGCTTTCACCAACTAAATTTGCCAGTATAAAT 1372  
Db |||||  
QY 413 TTGGTTAANATNAGATCCAAACTATCAGCTTTCNCAATTAANTGCGCCAGNTATA 354  
Db |||||  
QY 1373 AAAAGTAAGATAGGGCTTCTCAGCAGCAGCAGCAACTCCATCAGAAACTACTTTCAG 1432  
Db |||||  
QY 353 ATAAGGNAAGTGGTGTTCANGCAGCANCAGCAANTCCATCAGAAANTNCTTTCAG 294  
Db |||||  
QY 1433 CCGTCTACCAAAAAAGGAAAGGATGAAGAAATCAAGAAATGTCTTCATGCAATCA 1492  
Db |||||  
QY 293 CCGTNTACCAAAAAAGGAAAGGATGAAGAAATCAAGAAATGTNTTTCATGCAATCA 234  
Db |||||  
QY 1493 GCAAGATAGAAGCTGTGTTCTCTTTTAGAAACAAACCAACCTGCCTACCTCATTTG 1552  
Db |||||  
QY 233 GCAAGATAGAAGCTGTGTTCTTTTANAAACAAACCAACCTGNTACNCCCTCATTTG 174  
Db |||||  
QY 1553 TGGAAAAATAAGAGCAGCATCTATCTGAGAAATGAGCTGTGGACACAAACTCAGCAAT 1612  
Db |||||  
QY 173 TGGAAAAATAAGAGCAGCATNTATNTGAGATGAGCTGNGGACNCAAACTCAGNCAAT 114  
Db |||||  
QY 1613 AACTTATTACAGATACAGATTTAAATCTATTTGAAAAATCTGCCAGTAAATCTCAT 1672  
Db |||||  
QY 113 AACTTATTACAGATACAGATTTAAATNTATTTGTGAAAAATTTTGGCCAGTAAATNTCAT 54  
Db |||||

Qy 1673 GCTGCAGAAAAGCTTAGATCAATATAAAAAAGGGAAA 1709  
|||||  
Db 53 GCTGCAGAAAAGNTAGNTCCAAAAA 17

Search completed: August 30, 2005, 15:27:50  
Job time : 2563 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2005, 05:15:28 ; Search time 13109 Seconds  
(without alignments)  
12784.879 Million cell updates/sec

Title: US-09-837-138-1  
Perfect score: 4403  
Sequence: 1 ttcggcacaggcgcggttg.....accgcggtgagctccagct 4403

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_ges1:\*  
9: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4053.6	92.1	4692	3	HS0806925
2	4030.2	91.5	4550	3	BC040519 Homo sapi
3	1460.2	33.2	1497	3	BC005293 Homo sapi
4	1446.8	32.9	2265	9	AY416034 Homo sapi
5	1427	32.4	1460	3	BC016762 Homo sapi
6	1090.4	24.8	1792	3	BC031900 Mus muscu
7	1083	24.6	1731	3	AK031933 Mus muscu
8	913.2	20.7	964	5	BX405940
9	844.2	19.2	1021	5	BX405939
10	834.2	18.9	1163	4	BM542698
11	822.8	18.7	927	4	BG182890 RST1768 A
12	822.2	18.7	1803	9	AY416036 Mus muscu
13	771.4	17.5	785	7	CN304420
14	741.4	16.8	784	6	CB250418 UI-CF-FNO
15	741	16.8	789	4	BG214621 RST34264
16	734.8	16.7	747	7	CN304425
17	732.8	16.6	742	7	CN304424
18	732.6	16.6	1036	5	BX422568
19	732.4	16.6	902	4	BM461758
20	731.8	16.6	932	5	BU166634
21	725.4	16.5	775	4	BG676775
22	717.4	16.3	2285	9	AY416035 Pan trogl
23	715.8	16.3	771	5	BU620472 UI-H-FL1-
24	708.6	16.1	870	5	BU177917

25	708.2	16.1	818	7	CF593314
26	706.6	16.0	763	4	BM014420
27	702.8	16.0	855	5	BU517247
28	702.4	16.0	811	5	BQ574642 UI-H-EZ1-
29	696.2	15.8	960	4	BG284646
30	690.8	15.7	779	4	BG388866
31	689.4	15.7	691	6	CB123692
32	687.8	15.6	712	1	AI796269
33	676.2	15.4	735	4	BG779728
34	674	15.3	779	5	BU929703
35	673.2	15.3	707	5	BU661996
36	672.6	15.3	807	4	BG197194
37	671	15.2	752	1	AU18357
38	669.2	15.2	718	2	AW976050
39	664.4	15.1	747	6	CA777874
40	663	15.1	796	5	BM978914
41	663	15.1	845	4	BG532783
42	660.6	15.0	691	2	AW978306
43	660.6	15.0	751	7	CN304418
44	659.8	15.0	917	5	BP381003
45	651.2	14.8	706	4	BG292394

ALIGNMENTS

RESULT 1  
LOCUS HSM806925 4692 bp mRNA linear HTC 22-SEP-2004  
DEFINITION Homo sapiens mRNA; cDNA DKFP686G19151 (from clone DKFP686G19151).  
ACCESSION BX640816  
VERSION BX640816.1 GI:34365074  
KEYWORDS HTC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 4692)  
AUTHORS Bloembergen, H., Boecher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.  
CONSTRM The German CDNA Consortium  
TITLE Direct Submission  
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY  
COMMENT Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by GFP (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the CDNA sequencing consortium of the German Genome Project.  
This clone (DKFP686G19151) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:  
http://www.rzpd.de/cgi-bin/products/ci.cgi?CloneID=DKFP686G19151  
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES

Location/Qualifiers  
1..4692  
source  
/mol\_type="mRNA"  
/db\_xref="RZPD:DKFP686G19151Q"  
/db\_xref="taxon:9606"  
/clone="DKFP686G19151"  
/issue\_type="colon endothel, primary cell culture"  
/clone\_lib="686 (synonym: hicc3). Vector pSport1\_Sfi; host DH10B; sites SfiIA + SfiIB"  
/dev\_stage="adult"  
/note="Nijmegen breakage syndrome 1 (nibrin), differentially spliced, alternative start"  
join(407..1555,1557..2426)  
/gene="DKFP686G19151"  
join(407..1555,1557..2426)  
/gene="DKFP686G19151"  
/codon\_start=1

gene

CDS

/product="hypothetical protein"  
 /protein\_id="CAH56160.1"  
 /db\_xref="GI:52545912"  
 /translation="MQNGFSRTLKSGDGIITFVGFSKFRIEYELIVACSSLDIVSGKT  
 ALNQAILQGGFTVNNWTECHLVMVSVKVTIKTICALICGRPIIVKPEYTFEFLKAV  
 BSKQPPQIETSPVPLDEPSIGSKNDLSGQERKQIFKGTFFILNAKQHKLSAV  
 VFGGGEARLITENEBEHFFLAGPCVVDVTGITSQTLIPDCQKWIQSIMDLQRO  
 GLRPIPEAIGLAVIPMTTKNYCDPOGHPSTGLKTTTPGPSLSQGVSDKILMPAPV  
 NNTYVADTESQAUTDLSERPKEIKVSKMEQKFRMLSDQAPTVKESCKTSSNNNSM  
 VNTLIAMRIPNYQSPKLPINSKSRASQOQNTSIRNYFPQSTKRRDENDEKQ  
 MESSKARLETSCLEQTPATPSLWKNKEQHLSENPEVDTSNNLFTDIDLSKIV  
 KNSAKSHAAEKLRSNKGEMDDVAIEDEVLQLFKDTKPELEIDVKYQBEDVNDSE  
 KPRMDIETNDPDSBAVPESSKISOENIEIGKRRELKEDSLWSAKBISNNDKLODSE  
 MLPKLLITFEPSLVIKNSTSRNPSSINDDYQLKNFKFKVTVPGAKLPHIIGGS  
 DLIAHARXNTELEWLRQEMEVQNOHAKESLADDLFRYNFLKRRR"

## ORIGIN

Query Match 92.1%; Score 4053.6; DB 3; Length 4692;  
 Best Local Similarity 96.0%; Pred. No. 0;  
 Matches 4287; Conservative 9; Mismatches 28; Indels 142; Gaps 8;

QY 13 CGCGGTTGACGTCGCGCCAGCCCTGAGAGCCGACCGATGTGGAAACTGCTGCCGC 72  
 DB 71 CGCGGTTGACGTCGCGCCAGCCCTGAGAGCCGACCGATGTGGAAACTGCTGCCGC 130

QY 73 CGCGGCGCGGCGAGGAGAACCATACAGACTTTTGTAGCTGGGTTGAGTAGCTTTGG 132  
 DB 131 CGCGGCGCGGCGAGGAGAACCATACAGACTTTTGTAGCTGGGTTGAGTAGCTTTGG 190

QY 133 AAGGAAAACTGTGCCATTTCTAATTCAAAAATGATCAGTCGATCAGCCGAAATCATGCTGT 192  
 DB 191 AAGGAAAACTGTGCCATTTCTGATTTGAAATGATCAGTCGATCAGCCGAAATCATGCTGT 250

QY 193 GTTAACCTGTAACCTTTTGTAAACCAACCTG----- 223  
 DB 251 GTTAACCTGTAACCTTTCTGTAAACCAACCTGGTATAGAAAAACATTTTGGAAATTTCTAC 310

QY 224 -----AGTCAACAGATGAATCCCTGTTATGTGACATTAAGAT 262  
 DB 311 AGAAAAAGTAAATTTGTTACTAGTCAACAGATGAATCCCTGTTATGTGACATTAAGAT 370

QY 263 AATTCTAAGTATGGTACCTTTGTTAATGAGGAAAAAATGCAGAAATGGCTTTTCCGAACT 322  
 DB 371 AATTCTAAGTATGGTACCTTTGTTAATGAGGAAAAAATGCAGAAATGGCTTTTCCGAACT 430

QY 323 TTGAAGTCGGGGATGGTATTAATTTTGGAGTGTGTTGGAAGTAAATTCAGAAATAGAGTAT 382  
 DB 431 TTGAAGTCGGGGATGGTATTAATTTTGGAGTGTGTTGGAAGTAAATTCAGAAATAGAGTAT 490

QY 383 GAGCCTTTGGTGCATGCTCTCTGTTTGTAGTGTCTCTGGGAAACTGCTTTTAAATCAA 442  
 DB 491 GAGCCTTTGGTGCATGCTCTCTGTTTGTAGTGTCTCTGGGAAACTGCTTTTAAATCAA 550

QY 443 GCTATATGCAACTTCGAGGATTTACTGTAAACAAATGGACAGAAATGCCACTCACCTT 502  
 DB 551 GCTATATGCAACTTCGAGGATTTACTGTAAACAAATGGACAGAAATGCCACTCACCTT 610

QY 503 GTCATGGTATCAGTGAAGTTACCATTAACAAATATGTGCACTCAATTTGTGAACGTCCA 562  
 DB 611 GTCATGGTATCAGTGAAGTTACCATTAACAAATATGTGCACTCAATTTGTGAACGTCCA 670

QY 563 ATTGTAAGCCAGAAATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAAAGAGAGCCT 622  
 DB 671 ATTGTAAGCCAGAAATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAAAGAGAGCCT 730

QY 623 CCACAAATTTGAAAGTTTTTACCCACCTCTTGTATGAACCATCTATTGGAAGTAAAAATGTT 682  
 DB 731 CCACAAATTTGAAAGTTTTTACCCACCTCTTGTATGAACCATCTATTGGAAGTAAAAATGTT 790

QY 683 GATCTGTCAAGCGGAGGAAAGAAACAAATCTTCAAAGGGAAACATTTATATTTTTG 742  
 DB 791 GATCTGTCAAGCGGAGGAAAGAAACAAATCTTCAAAGGGAAACATTTATATTTTTG 850

743 AATGCCAAACAGCATAGAATAAGTTAGTTCCCGAGTGTGTCTTTGGAGGTGGGAAGCTAGG 802  
 DB 851 AATGCCAAACAGCATAGAATAAGTTAGTTCCCGAGTGTGTCTTTGGAGGTGGGAAGCTAGG 910

QY 803 TTGTATAACAGAGAAGTAAAGAGAAACATTAATTTCTTTTGGTCCCGGGAACGTGTGTT 862  
 DB 911 TTGTATAACAGAGAAGTAAAGAGAAACATTAATTTCTTTTGGTCCCGGGAACGTGTGTT 970

QY 863 GTTGATACAGGAATAACAACTCAGAGACCTTAATCTCTGATCTGTCAAGAAATGGATT 922  
 DB 971 GTTGATACAGGAATAACAACTCAGAGACCTTAATTTCTCTGATCTGTCAAGAAATGGATT 1030

QY 923 CAGTCAATAATGAGTATGCTCTCAAAGGCAAGGTCTTAGACCTATTCTCTGAAGCAGAAAT 982  
 DB 1031 CAGTCAATAATGAGTATGCTCTCAAAGGCAAGGTCTTAGACCTATTCTCTGAAGCAGAAAT 1090

QY 983 GGATTTGGCGGTGATTTTCATGACTACAAGAAATTAATCTGTGATCTCTCAGGGCCATCCCCAGT 1042  
 DB 1091 GGATTTGGCGGTGATTTTCATGACTACAAGAAATTAATCTGTGATCTCTCAGGGCCATCCCCAGT 1150

QY 1043 ACAGGATTAAGACAACTCCAGGACCAAGCCCTTTCAAGGCGGTGTCAAGTATGATGAA 1102  
 DB 1151 ACAGGATTAAGACAACTCCAGGACCAAGCCCTTTCAAGGCGGTGTCAAGTATGATGAA 1210

QY 1103 AAACTAATGCCAAGCGCCCTCAGTGAACACTACAACTACGTAGCTGACACAGAAATCAGAG 1162  
 DB 1211 AAACTAATGCCAAGCGCCCTCAGTGAACACTACAACTACGTAGCTGACACAGAAATCAGAG 1270

QY 1163 CAAGCAGATACATGGGATTTGAGTGAAGGCAAGAAATCAAAAGTCTCCAAAATGGAA 1222  
 DB 1271 CAAGCAGATACATGGGATTTGAGTGAAGGCAAGAAATCAAAAGTCTCCAAAATGGAA 1330

QY 1223 CAAAATTTCAAGATGCTTTTCAAGACCGACCCACTGTAAAGGAGTCTGCAAAACAGC 1282  
 DB 1331 CAAAATTTCAAGATGCTTTTCAAGACCGACCCACTGTAAAGGAGTCTGCAAAACAGC 1390

QY 1283 TCTAATAATAATAGTATGATGATCAAAATCTTTGGCTAAGATGAGAAATCCCAAACTATCAG 1342  
 DB 1391 TCTAATAATAATAGTATGATGATCAAAATCTTTGGCTAAGATGAGAAATCCCAAACTATCAG 1450

QY 1343 CTTTCCAACTAAATTTGCCAAGTATAAATAAAGTAAAGATAGGGCTTCTCAGCAGCAG 1402  
 DB 1451 CTTTCCAACTAAATTTGCCAAGTATAAATAAAGTAAAGATAGGGCTTCTCAGCAGCAG 1510

QY 1403 CAGACCAACTCCATCAGAACTTCTTTTCCAGCCGCTACCC - AAAAAGGGAAGGAGTGA 1461  
 DB 1511 CAGACCAACTCCATCAGAACTTCTTTTCCAGCCGCTACCCAAAAGGGAAGGAGTGA 1570

QY 1462 AGAAATCAAGAAATGTCTTCATGCAAAATCAGCAAGAAATAGAAACGTCTTGTCTCTTTT 1521  
 DB 1571 AGAAATCAAGAAATGTCTTCATGCAAAATCAGCAAGAAATAGAAACGTCTTGTCTCTTTT 1630

QY 1522 AGAACAAACACAACTGCTTACCCCTCATTTGTGAAAAAATAAGGAGCAGCATCTATCTGA 1581  
 DB 1631 AGAACAAACACAACTGCTTACCCCTCATTTGTGAAAAAATAAGGAGCAGCATCTATCTGA 1690

QY 1582 GAATGAGCCTGTGGACACAACTCAGACAACTTATTTTACAGATCAGATTTTAAATC 1641  
 DB 1691 GAATGAGCCTGTGGACACAACTCAGACAACTTATTTTACAGATCAGATTTTAAATC 1750

QY 1642 TATTGTGAAAAATTTCTGCGCAGTAAATCTCTGCTCAGAAAAAGCTTAAGATCAAAATAAAA 1701  
 DB 1751 TATTGTGAAAAATTTCTGCGCAGTAAATCTCTGCTCAGAAAAAGCTTAAGATCAAAATAAAA 1810

QY 1702 AAGGAAATGAGATGATGTGGCCATAGAAAGATGAAATTTGGAAACAGTTATTCAAGACAC 1761  
 DB 1811 AAGGAAATGAGATGATGTGGCCATAGAAAGATGAAATTTGGAAACAGTTATTCAAGACAC 1870

QY 1762 AAAACAGAGTTAGAAATTTGATGAAAGTTCAAAACAGGAGGAGATGCAATGTTAG 1821  
 DB 1871 AAAACAGAGTTAGAAATTTGATGAAAGTTCAAAACAGGAGGAGATGCAATGTTAG 1930

QY 1822 AAAAGGCCAAGGATGGATATAGAAACAAATGACACTTTTCAAGTATGAGCAGTACCAGA 1881

|||||  
Db 1931 AAAAAAGCCAGGATGGATATAGAAAACAAATGACACTTTTCAGTGATGAAGCAGTACCAGA 1990  
QY 1882 AAGTAGCAAAATATCTCAAGAAAATGAAATTTGGGAAGAAACGTGAACCTCAAGGAAGACTC 1941  
Db 1991 AAGTAGCAAAATATCTCAAGAAAATGAAATTTGGGAAGAAACGTGAACCTCAAGGAAGACTC 2050  
QY 1942 ACTATGGTCAGCTAAAGAAATATCTAAACAATGACAACTTCAGGATGATAGTAGATGCT 2001  
Db 2051 ACTATGGTCAGCTAAAGAAATATCTAAACAATGACAACTTCAGGATGATAGTAGATGCT 2110  
QY 2002 TCCAAAAAGCTGTTATTCAGCTCAATTTAGATCACTGGTGATTAAGAACTCTACTTCCAG 2061  
Db 2111 TCCAAAAAGCTGTTATTCAGCTCAATTTAGATCACTGGTGATTAAGAACTCTACTTCCAG 2170  
QY 2062 AATCCGCTCGGCATAAATGATGATTTATGCTCAACTAAAGAAATTTCAAGAAATTTCAAAA 2121  
Db 2171 AATCCATCTGGCATAAATGATGATTTATGCTCAACTAAAGAAATTTCAAGAAATTTCAAAA 2230  
QY 2122 GGTCAATATCTGGAGCAGGAAAACTTCCACACATCATTTGGAGGATCAGATCTAATAGC 2181  
Db 2231 GGTCAATATCTGGAGCAGGAAAACTTCCACACATCATTTGGAGGATCAGATCTAATAGC 2290  
QY 2182 TCATCATGCTCGAAAGAAATACAGAACTAGAGAGTGGCTAAGCGCAGGAAATGGAGTACA 2241  
Db 2291 TCATCATGCTCGAAAGAAATACAGAACTAGAGAGTGGCTAAGCGCAGGAAATGGAGTACA 2350  
QY 2242 AATCAACATGCAAAAGAGAGTCTCTTCTGATGATCTTTTTPAGATACAACTCTTATTT 2301  
Db 2351 AATCAACATGCAAAAGAGAGTCTCTTCTGATGATCTTTTTPAGATACAACTCTTATTT 2410  
QY 2302 AAAAAAGGAGAAGATACTGAGGATTTTAAAAAGAACCCATGGAAAACTTCTCTAGTAAGC 2361  
Db 2411 AAAAAAGGAGAAGATACTGAGGATTTTAAAAAGAACCCATGGAAAACTTCTCTAGTAAGC 2470  
QY 2362 ATCTACTTCAGGCCACACAGGTTATATGATATATATATATATATATATATATATATATAT 2421  
Db 2471 ATCTACTTCAGGCCACACAGGTTATATGATATATATATATATATATATATATATATAT 2530  
QY 2422 AATGTTTTATGGCTAAATTTTAAATAAATGCAAAAACTTTGATTTCTTTTGTATGT 2481  
Db 2531 AATGTTTTATGGCTAAATTTTAAATAAATGCAAAAACTTTGATTTCTTTTGTATGT 2590  
QY 2482 AACAAATGTTGTGTCGTTTTTCAGGCTTTGTCAATGCTCATCTTTTTCATTTTAAATGT 2541  
Db 2591 AACAAATGTTGTGTCGTTTTTCAGGCTTTGTCAATGCTCATCTTTTTCATTTTAAATGT 2650  
QY 2542 GTTTTGTTTTAAATAGTTAATATATAGTACAGTTCAAAATTTCTAAATATACGTAAAGGTA 2601  
Db 2651 GTTTTGTTTTAAATAGTTAATATATAGTACAGTTCAAAATTTCTAAATATACGTAAAGGTA 2710  
QY 2602 AAGGACTTAAAGTCACCCCTTCCACCATTTGCTCCTAGCTACT- - - - - 2640  
Db 2711 AA-GACTTAAAGTCACCCCTTCCACCATTTGCTCCTAGCTACTTGGTTCCTCCCTCAGNAAAAAT 2769  
QY 2641 - - - - - 2642  
Db 2770 TCATGATACTCTTTCTTATGAATCTTTCCAGGATTTTGTAGTCCCTATTTCAAAATTCCTTA 2829  
QY 2643 TTTTAAATTAATTTCTTACAAATGATAGCATAAATATGATGAGTGTCTTACACCTTGCT 2702  
Db 2830 TTTTAAATTAATTTCTTACAAATGATAGCATAAATATGATGAGTGTCTTACACCTTGCT 2889  
QY 2703 TTTTAACTTAGTAAGATTTAAATTAATAGGAATATCAATATATATATATATATATATTTT 2762  
Db 2890 TTTTAACTTAGT-AGATTAATTAATTAATAGGAATATCAATATATATATATATATTTT 2948  
QY 2763 CTTTTCATATGCTGTAGTCTTACCTAAACTCTGGTATCCAAACAAATGGCTTCAGT 2822  
Db 2949 CTTTTCATATGCTGTAGTCTTACCTAAACTCTGGTATCCAAACAAATGGCTTCAGT 3008  
QY 2823 GGTGAGATGTCACCTACATGTTATTTCTAGTACTAGAAATCTGAGAACCAATGTGGAGACTT 2882  
|||||

Db 3009 GGTGAGATGTCACCTACATGATTTATTTAGTACTAGAACTGAGAACCAATGTGGAGACTT 3068  
QY 2883 CATCAAAACATGGGTTTAGTTTTTCCACAGAAATGAAAGACCTGTACCCCTTTTTTGGTGGTC 2942  
Db 3069 CATCAAAACATGGGTTTAGTTTTTCCACAGAAATGAAAGACCTGTACCCCTTTTTTGGTGGTC 3128  
QY 2943 TTACTGAGCTGGGTGGGTGTCGTTTTTGTAGCTTATTTAGAGTCTTAGTCTTCTACTTAT 3002  
Db 3129 TTACTGAGCTGGGTGGGTGTCGTTTTTGTAGCTTATTTAGAGTCTTAGTCTTCTACTTAT 3188  
QY 3003 AAGTAGAAATCGTGAGATTTGTTTTTCTTCTTACCTAAAGGGAGATGTAAGAAACAA 3062  
Db 3189 AAGTAGAAATCGTGAGATTTGTTTTTCTTCTTACCTTAAAGGGAGATGTAAGAAACAA 3248  
QY 3063 TGAATGTCCTTTTTTCAAACTTTATTTGACAAGTGATTTTCAAGTCTGTGTTCAAAAATATA 3122  
Db 3249 TGAATGTCCTTTTTTCAAACTTTATTTGACAAGTGATTTTCAAGTCTGTGTTCAAAAATATA 3308  
QY 3123 TTCATGTACTCTGTATCCAGCAAGAGGGAGTTCCAGTCAAGAGTCACTACAACTGATTA 3182  
Db 3309 TTCTATGTACTCTGTATCCAGCAAGAGGGAGTTCCAGTCAAGAGTCACTACAACTGATTA 3368  
QY 3183 GTTGTTTTAGAGAAATGAGAAATGGAACAGTGAGGAATGGAGGCCATATTTCCATGACTTCC 3242  
Db 3369 GTTGTTTTAGAGAAATGAGAAATGGAACAGTGAGGAATGGAGGCCATATTTCCATGACTTCC 3428  
QY 3243 CTTTAAACAGAAAGCAACAGAAAGGAGGCTGGCTCTTACATCACTCTCACCTTCC 3302  
Db 3429 CTTGTAAACAGAAAGCAACAGAAAGGAGCAAGAGGCTGGCTCTTACATCACTCTCACCTTCC 3488  
QY 3303 AATCTTGTGGAAGTGATCTACTTGCAGAAACCAAAATTAACCTTCTCAAGTCTGGC 3362  
Db 3489 AATCTTGTGGAAGTGATCTACTTGCAGAAACCAAAATTAACCTTCTCAAGTCTGGC 3548  
QY 3363 TCGTTGCGAGGTGGAACCTCCAGCTGCAAGGAGGTAGGGAATGGAAGTCTTTTTTAA 3422  
Db 3549 TCGTTGCGAGGTGGAACCTCCAGCTGCAAGGAGGTAGGGAATGGAAGTCTTTTTTAA 3608  
QY 3423 GCTTCTCAGCCTTCTTAGGGAACAGAAATTTGGGTGAGGCAATCTGCAATTTCTACTACAG 3482  
Db 3609 GCTTCTCAGCCTTCTTAGGGAACAGAAATTTGGGTGAGGCAATCTGCAATTTCTACTACAG 3668  
QY 3483 GCAATTGAGACAGTGTAGATTAATTAATAGAGATTAAGAACACTTAAATATGA 3542  
Db 3669 GCAATTGAGACAGTGTAGATTAATTAATAGAGATTAAGAACACTTAAATATGA 3728  
QY 3543 TAGTGGTATGACATTTGGATAGAACCTGGATCTTTAGAGTAGAATTTGACAGGCGATAT 3602  
Db 3729 TAGTGGTATGACATTTGGATAGAACCTGGATCTTTAGAGTAGAATTTGACAGGCGATAT 3788  
QY 3603 TAGTTGATGAATGGAGTCAATTTGAGTCTTTPAATAGCCATGTATCATAAATTTACCAAGTG 3662  
Db 3789 TAGTTGATGAATGGAGTCAATTTGAGTCTTTPAATAGCCATGTATCATAAATTTACCAAGTG 3848  
QY 3663 AAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATTAATGGACAGATTAATA 3722  
Db 3849 AAGCTGGTGGAAACATATGGTCTCCATTTTACAGTTAAGGAATATTAATGGACAGATTAATA 3908  
QY 3723 TTGTTTCTGTGTCATGCCCACTCCCTTTCTAAGGAAGACTGCCCTACTATAGCAGTTTTT 3782  
Db 3909 TTGTTTCTGTGTCATGCCCACTCCCTTTCTAAGGAAGACTGCCCTACTATAGCAGTTTTT 3968  
QY 3783 TATATTTGTCAATTTTATGAATATGAATGAGGAGTTCTGTGATCCTCTCTCTTTTACAA 3842  
Db 3969 TATATTTGTCAATTTTATGAATATGAATGA- GAGTTCTGGTACTCTCTCTCTTTACAA 4027  
QY 3843 ATATTGGGTGTCAGATTTTCCCTTTTAAACCMWTCCAAATTCGGGTGTGTAGGT 3902  
Db 4028 ATATTGG- - - - - TGTGTGTCAGTATTTTCTTTTAAACCAATTTCAATTCGGGTGTGTAGGT 4083  
QY 3903 GATATTGTTCCATTTGGGTTTTTAAATTTGATATATCTCTGATAGCTATATTTGGGTCATAGAA 3962  
Db 4084 G- - - - - TTTTCATTTTGGTTTTTAAATTTGATATATCCCTGATAGCTATTAATTTGGGTCATAGAA 4138  
|||||

3963 ATTCTTTATACATCTTAGATGCAAGTCTCTTGGCGATATACGTATTGTGATATTACACC 4022  
 Db ATTCCTTTATACATCTTAGATGCAAGTCTCTTGGCGATATATGTATTGTGATATTACACC 4198  
 QY TAGCTGTGGCTGACTGTTTCTTTATGTCCTTTGATGAATAGAAAGTTTAAATTTTGA 4082  
 Db TAGCTGTGGCTGACTGTTTCTTTATGTCCTTTGATGAATAGAAAGTTTAAATTTTGA 4258  
 QY CAAGGTCAAATTTATTTTCTTTTGTGATATTTTCTCTCAATTTAAACCCCAAG 4142  
 Db CAAGGTCAAATTTATTTTCTTTTGTGATATTTTCTCTCAATTTAAACCCCAAG 4318  
 QY ATTTTCAGATATCTGCTCTATATATAAATTTTATATTTTATTTGTGATCTACCTTG 4202  
 Db ATTTTCAGATATCTGCTCTATATATAAATTTTATATTTTATTTGTGATCTACCTTG 4378  
 QY AATTGATATGATGTTGTGAATATGATGACAGGGTCTTTTTCCTCCCATCAAGATATC 4262  
 Db AATTGATATGATGTTGTGAATATGATGACAGGGTCTTTTTCCTCCCATCAAGATATC 4438  
 QY CAGTCATTTGACATCTGTTTATGAAAGATATATCTTCTTCATTAATTAACCTTGCCA 4322  
 Db CAGTCATTTGACATCTGTTTATGAAAGATATATCTTCTTCATTAATTAACCTTGCCA 4498  
 QY ATTAGTAAAAAATCAATTAACCATM 4348  
 Db ATTAGTAAAAAATCAATTAACCATAA 4524

## RESULT 2

BC040519 4550 bp mRNA linear HTC 19-NOV-2003  
 LOCUS Homo sapiens Nijmegen breakage syndrome 1 (nibrin), mRNA (cdna  
 DEFINITION clone IMAGE:4791579), containing frame-shift errors.

ACCESSION BC040519

VERSION BC040519.1 GI:26996492

KEYWORDS HTC

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE 1 (bases 1 to 4550)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,  
 Schetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S.,  
 Carninci P., Frange C., Raha S.S., Loquellano N.A., Peters G.J.,  
 Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J.,  
 McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,  
 Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahey J., Helton E., Kettman M., Madan A., Rodrigues S.,  
 Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y.,  
 Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,  
 Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 Butterfield Y.S., Krzywinski M.I., Skalski U., Smalls D.E.,  
 Scherch A., Schein J.E., Jones S.J. and Marra M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 4550)

Strausberg R.

Direct Submission

Submitted (29-NOV-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

## REMARK

## COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAP Plate: 63 Row: a Column: 1  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: matched mRNA gi: 6996019  
 This clone has the following problem: frame shifted.

## FEATURES

## source

1..4550  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4791579"  
 /tissue\_type="Brain, hypothalamus"  
 /clone\_lib="NIH\_MGC\_96"  
 /lab\_host="DH10B"  
 /note="Vector: pBluescript"

## ORIGIN

Query Match 91.5%; Score 4030.2; DB 3; Length 4550;  
 Best Local Similarity 95.6%; Pred. No. 0;  
 Matches 4277; Conservative 16; Mismatches 34; Indels 146; Gaps 9;  
 QY 13 CGCGGTTGCACGTCGGCCCGCCCGCCGATGAGGAGCGGCGGATGTGGAACTGCTCCCGC 72  
 Db 71 CGCGGTTGCACGTCGGCCCGCCCGCCGATGAGGAGCGGCGGATGTGGAACTGCTCCCGC 130  
 QY 73 CGCGGCGCGGCGAGGAGAACCATACAGACTTTTGTGACTGGCGTTGAGTGTGTTGG 132  
 Db 131 CGCGGCGCGGCGAGGAGAACCATACAGACTTTTGTGACTGGCGTTGAGTGTGTTGG 190  
 QY 133 AAGGAAATCTGTCCTTCTTAATGAAATGATCAGTCGATCAGCCGCAATCATGCTGT 192  
 Db 191 AAGGAAATCTGTCCTTCTTAATGAAATGATCAGTCGATCAGCCGCAATCATGCTGT 250  
 QY 193 GTTAATCTGCTAACTTTTCTGTAACCAACTG----- 223  
 Db 251 GTTAATCTGCTAACTTTTCTGTAACCAACTGTTATAGAAAACATTTTGGAAATTTCTAC 310  
 QY 224 -----AGTCAACACAGATGAAATCCCTGTATTGACATTTAAAGAT 262  
 Db 311 AGAAAAAGTAAATTTGTTACTAGTCAACACAGATGAAATCCCTGTATTGACATTTAAAGAT 370  
 QY 263 AATTTCTAGTATGTCCTTTGTTAATCAGAAAATAATGCAAGTGGCTTTTCCGAACT 322  
 Db 371 AATTTCTAGTATGTCCTTTGTTAATCAGAAAATAATGCAAGTGGCTTTTCCGAACT 430  
 QY 323 TTGAAGTCGGGGATGGTATTACTTTTGGAGTGTGTTGAAAGTAAATTCAGAAATAGATAT 382  
 Db 431 TTGAAGTCGGGGATGGTATTACTTTTGGAGTGTGTTGAAAGTAAATTCAGAAATAGATAT 490  
 QY 383 GAGCCTTTGGTGTGATGCTCTTCTTTTGTAGATGCTCTGGGAAATCTGCTTTTAAATCAA 442  
 Db 491 GAGCCTTTGGTGTGATGCTCTTCTTTTGTAGATGCTCTGGGAAATCTGCTTTTAAATCAA 550  
 QY 443 GCTATTATGCACTTGGAGGATTTTACTTAAACATTCGACAGAGATGCACTCACCTT 502  
 Db 551 GCTATTATGCACTTGGAGGATTTTACTGTAACAATTTGACAGAGATGCACTCACCTT 610  
 QY 503 GTCATGGTATCAGTGAAAGTTTACCATTAAAAACAATATATGCACTCATTTTGTGGACGTCCA 562

Db 611 GTATGGTATCAGTGAAGTTACATTTAAACCAATATGTGCACTCATTTGTGACGCTCCA 670  
Qy 563 ATTGTAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTCAGTCCAAAGAACGAGCCT 622  
Db 671 ATTGTAAGCCAGAAATATTTTACTGAAATTCCTGAAAGCAGTTCAGTCCAAAGAACGAGCCT 730  
Qy 623 CCAAAATGAAGTTTTTACCACCTCTTTGATGAACCATCTATTTGGAAGTAAAAATGTT 682  
Db 731 CCAAAATGAAGTTTTTACCACCTCTTTGATGAACCATCTATTTGGAAGTAAAAATGTT 790  
Qy 683 GATCTGTGAGCGGAGGAAGAAACAATCTTCAAGGGGAAACATTTATATTTTTC 742  
Db 791 GATCTGTGAGCGGAGGAAGAAACAATCTTCAAGGGGAAACATTTATATTTTTC 850  
Qy 743 AATGCCAAAACAGCATAGAATTTGAGTTCGCGAGTTGTCTTTGAGAGTGGGGAAGCTAGG 802  
Db 851 AATGCCAAAACAGCATAGAATTTGAGTTCGCGAGTTGTCTTTGAGAGTGGGGAAGCTAGG 910  
Qy 803 TTGATAACAGAGAGAAATGAAGAAACAATAATTTCTTTTGGCTCCGGGAACGTTGTT 862  
Db 911 TTGATAACAGAGAGAAATGAAGAAACAATAATTTCTTTTGGCTCCGGGAACGTTGTT 970  
Qy 863 GTTGATACAGGAATACAAATCACAGACCTTAATTCCTGACTGTGAGAGAAATGGAAT 922  
Db 971 GTTGATACAGGAATACAAATCACAGACCTTAATTCCTGACTGTGAGAGAAATGGAAT 1030  
Qy 923 CAGTCAATATGATATGCTCCAAAGCAAGGCTTTAGACCTAATTCCTGAAGCAGAAAT 982  
Db 1031 CAGTCAATATGATATGCTCCAAAGCAAGGCTTTAGACCTAATTCCTGAAGC -- AAT 1088  
Qy 983 GGATGGCGTGATTTTTCATGACTACAAAGAAATTAAGTGTGATCCTCAGGSCCATCCAGT 1042  
Db 1089 GGATGGCGTGATTTTTCATGACTACAAAGAAATTAAGTGTGATCCTCAGGSCCATCCAGT 1148  
Qy 1043 ACAGGATTAAGAACAACTCCAGGACCAAGCCTTTTCAAGCGGTGTCAGTTGATGA 1102  
Db 1149 ACAGGATTAAGAACAACTCCAGGACCAAGCCTTTTCAAGCGGTGTCAGTTGATGA 1208  
Qy 1103 AAATAATGCAAGCGCCAGTGAAACACTACAAACATACGTAGCTGACACAGAAATCAGAG 1162  
Db 1209 AAATAATGCAAGCGCCAGTGAAACACTACAAACATACGTAGCTGACACAGAAATCAGAG 1268  
Qy 1163 CAAGCAGATACATGGGATTTGAGTGAAGGCCAAAAGAAATCAAGTCTCCAAATGGAA 1222  
Db 1269 CAAGCAGATACATGGGATTTGAGTGAAGGCCAAAAGAAATCAAGTCTCCAAATGGAA 1328  
Qy 1223 CAAAAATTCAGATGCTTTTCAAGAGCGCACCTGTTAAAGAGTCTTGCAAAAACAGC 1282  
Db 1329 CAAAAATTCAGATGCTTTTCAAGAGTGCACCCACTGTAAAGAGTCTTGCAAAAACAGC 1388  
Qy 1283 TCTAATAATAATAGTATGTTATCAAAATCTTTTGGCTAAGATGAGAAATCCCAAACTATCAG 1342  
Db 1389 TCTAATAATAATAGTATGTTATCAAAATCTTTTGGCTAAGATGAGAAATCCCAAACTATCAG 1448  
Qy 1343 CTTTCCACCAACTAAATGGCAAGTATAAATAAAGTAAAGATAGAGTGGGCTTCTCAGCAGCAG 1402  
Db 1449 CTTTCCACCAACTAAATGGCAAGTATAAATAAAGTAAAGATAGAGTGGGCTTCTCAGCAGCAG 1508  
Qy 1403 CAGACCAACTCCATCAGAACTACTTTAGCCCTCTACCAAAAAGGGAAGGAGTGA 1462  
Db 1509 CAGACCAACTCCATCAGAACTACTTTAGCCCTCTACCAAAAAGGGAAGGAGTGA 1568  
Qy 1463 GAAATTCAGAAATGCTTCATGCAATCAGCAAGAAATGAAACGCTTGTCTCTTTT 1522  
Db 1569 GAAATTCAGAAATGCTTCATGCAATCAGCAAGAAATGAAACGCTTGTCTCTTTT 1628  
Qy 1523 GAAACAAACAACTGCTACACCTCATTTGTGAAAAATAAGGAGCAGCATCTATCTGAG 1582  
Db 1629 GAAACAAACAACTGCTACACCTCATTTGTGAAAAATA -- AGCAGCATCTATCTGAG 1685  
Qy 1583 AATGAGCCTGTGAGACAAACTCAGACAAATACTTATTAAGATACAGATACAGATTTAAATCT 1642

Db 1686 AATGAGCCTGTGGACACAAAATCAGACAATAACTTATTTACAGATACAGATTTAAATCT 1745  
Qy 1643 ATTGTGAAAAATTTCTGCCAGTAAATCTCATGTGCAAGAAAAGCTAGATCAAAATAAAAA 1702  
Db 1746 ATTGTGAAAAATTTCTGCCAGTAAATCTCATGTGCAAGAAAAGCTAGATCAAAATAAAAA 1805  
Qy 1703 AGGGAATGGATGATGTGGCCATAGAAGATGAAGTATTGGAAACAGTATTCAAGGACACA 1762  
Db 1806 AGGGAATGGATGATGTGGCCATAGAAGATGAAGTATTGGAAACAGTATTCAAGGACACA 1865  
Qy 1763 AAACACAGCTTAGAAATTTGATGTGAAAGTTCAAAACAGGAGGAGATGCTCAATGTTAGA 1822  
Db 1866 AAACACAGCTTAGAAATTTGATGTGAAAGTTCAAAACAGGAGGAGATGCTCAATGTTAGA 1925  
Qy 1823 AAAAGCCCAAGGATGGATATAGAAAACAAATGACACTTTTCAGTGTAGAACAGTACCAGAA 1882  
Db 1926 AAAAGCCCAAGGATGGATATAGAAAACAAATGACACTTTTCAGTGTAGAACAGTACCAGAA 1985  
Qy 1883 AGTAGCAAAATATCTCAAGAAAATGAAATTTGGGAAGAAACGTTGAACCTCAAGGAGACTCA 1942  
Db 1986 AGTAGCAAAATATCTCAAGAAAATGAAATTTGGGAAGAAACGTTGAACCTCAAGGAGACTCA 2045  
Qy 1943 CTATGTCAGCTAAAGAAATATCTACCAATGACAACTTTTCAGATGATAGTGAGATGCTT 2002  
Db 2046 CTATGTCAGCTAAAGAAATATCTACCAATGACAACTTTTCAGATGATAGTGAGATGCTT 2105  
Qy 2003 CCAAAAAAGCTGTTATTTGACTGAAATTTAGATCACTGGTGATTTAAAAAATCTTACTTCCAGA 2062  
Db 2106 CCAAAAAAGCTGTTATTTGACTGAAATTTAGATCACTGGTGATTTAAAAAATCTTACTTCCAGA 2165  
Qy 2063 AATCGTCTGGCATAAATGATGATTTATGTTCACTCACTAAAAAATTTCAAGAAAATTTCAAAAAG 2122  
Db 2166 AATCCATCTGGCATAAATGATGATTTATGTTCACTCACTAAAAAATTTCAAGAAAATTTCAAAAAG 2225  
Qy 2123 GTCAATATCTGGAGCAGGAAAATCTTCCACATCATTTGGAGGATCAGATCTTAATAGCT 2182  
Db 2226 GTCAATATCTGGAGCAGGAAAATCTTCCACATCATTTGGAGGATCAGATCTTAATAGCT 2285  
Qy 2183 CATCATGCTCGAAAGAAACAGAACTAGAAAGTGGCTAAGCAGGAAATGGAGGTACAA 2242  
Db 2286 CATCATGCTCGAAAGAAACAGAACTAGAAAGTGGCTAAGCAGGAAATGGAGGTACAA 2345  
Qy 2243 AATCAACATGCAAAAAGAGAGTCTTGTGTGATGATCTTTTGTAGATACAACTCTTATTTA 2302  
Db 2346 AATCAACATGCAAAAAGAGAGTCTTGTGTGATGATCTTTTGTAGATACAACTCTTATTTA 2405  
Qy 2303 AAAGGAGAGATTAATCTGAGGATTTTAAAGAACCCATGGAAACCTTCCCTAGTAAGCA 2362  
Db 2406 AAAGGAGAGATTAATCTGAGGATTTTAAAGAACCCATGGAAACCTTCCCTAGTAAGCA 2465  
Qy 2363 TCTACTTCAGGCAACAGGTTTATATGAATATATAGTATATAGAACGATTTTAAAGTTACA 2422  
Db 2466 TCTACTTCAGGCAACAGGTTTATATGAATATATAGTATATAGAACGATTTTAAAGTTACA 2525  
Qy 2423 ATGTTTTATGGCTAAAATTTAATAAATGCAAAAACTTTGATCTTTTGTATGTA 2482  
Db 2526 ATGTTTTATGGCTAAAATTTAATAAATGCAAAAACTTTGATCTTTTGTATGTA 2585  
Qy 2483 ACAAATGTTGTTCTGTTTTTCAGGCTTTGTCAATGCATCTTTTTCATTTTTTAAATG 2542  
Db 2586 ACAAATGTTGTTCTGTTTTTCAGGCTTTGTCAATGCATCTTTTTCATTTTTTAAATG 2645  
Qy 2543 TTTTGTTTTAAATAGTTAATATAGTACAGTTCAAAATTTCTAAATRTACGTAAGGTAA 2602  
Db 2646 TTTTGTTTTAAATAGTTAATATAGTACAGTTCAAAATTTCTAAATRTACGTAAGGTAA 2705  
Qy 2603 AGGACTAAAGTCACTTCCACCATGCTAGTACT----- 2640  
Db 2706 A-GACTAAAGTCACTTCCACCATGCTAGTACTTGGTTCCTCCCTCAGAAAAAAAT 2764  
Qy 2641 -----TAT 2643  
Db 2765 CATGATACTCATTTCTTATGAATCTTTCCAGGATTTTGTGAGTCTTATTCAAATTCCTAT 2824

QY	2644	TTTTAATAATTTCCCTACACAAATGATAGACATATGCGAGTGTTCTACACCTTGGCTT	2703	QY	3724	TGTTTCTGTGCATGCCACCAATCCCTTCTTAAGGAAGACTGCCCTACTATAGACGTTTTT	3783
Db	2825	TTTTAAATATTTTCCCTACACAAATGATAGACATATGCGAGTGTTCTACACCTTGGCTT	2884	Db	3904	TGTTTCTGTGCATGCCACCAATCCCTTCTTAAGGAAGACTGCCCTACTATAGACGTTTTT	3963
QY	2704	TTTTTACTTAGTAGAATTAATAAATATATAGGAATATCAATATAATGTTTTTAATATTTTTTC	2763	QY	3784	ATATTGTCAATTTATGAATATAATGAATGAGGAGTCTGGTACCTCTGTCTTTTACAAA	3843
Db	2885	TTTTTACTTAGT-AGATTAAATAATAGGAATATCAATATAATGTTTTTAATATTTTTTC	2943	Db	3964	ATATTGTCAATTTATGAATATAATGAATGA- GAGTTCCTGGTACCTCTGTCTTTTACAAA	4022
QY	2764	TTTTTCCATTATGCTGTAGTCTTACCTAAACTCTGGTGATCCCAACAAATGGCTTCAGTG	2823	QY	3844	TATTGGGTGGTGCAGTATTTTTTCCCTTTTAAACCMTCCCAATTCGGGTGGTGTAGTG	3903
Db	2944	TTTTTCCATTATGCTGTACTCTTACCTAAACTCTGGTGATCCCAACAAATGGCTTCAGTG	3003	Db	4023	TATTGG- - - - -TGTTGTCAGTATTTTTTCCCTTTTAAACCAATCCAAATCGGTGGTGTAGTGATG	4078
QY	2824	GTGCGAGTGTCACTACATGTTATTTCTAGTACTAGAACTGAGACCAATGGGAGACTTC	2883	QY	3904	GATGTTTTCCATTTGGGTTTTTAATTTGTATATCCCTGATAGCTATATAATTTGGGTGATAGAAA	3963
Db	3004	GTGCGAGTGTCACTACATGTTATTTCTAGTACTAGAACTGAGACCAATGGGAGACTTC	3063	Db	4079	-----TTTTCATTTTGGTTTTTAATTTGTATATCCCTGATAGCTATATAATTTGGGTGATAGAAA	4133
QY	2884	ATCAAAATGGGTTTTAGTTTTTACACAGAAATGGAAGACCTGTACCCCTTTTTTGGTGGTCT	2943	QY	3964	TTCTTTTATACATTTAGATGCAAGTCTCTTGTGCGATATAGCTATTTGAGATATTTACACT	4023
Db	3064	ATCAAAATGGGTTTTAGTTTTTACACAGAAATGGAAGACCTGTACCCCTTTTTTGGTGGTCT	3123	Db	4134	TTCTTTTATACATTTAGATGCAAGTCTCTTGTGCGATATAGCTATTTGAGATATTTACACT	4193
QY	2944	TACTGAGCTGGGTGGTGTCTGTTTTGAGCTTATTTAGAGTCTAGTTTTCTCTACTTATA	3003	QY	4024	AGTCTGTGGCTTGACTGTTTTTCTTATGCTTTTGGATGAATAGAGTTTTTAAATTTTGAC	4083
Db	3124	TACTGAGCTGGGTGGTGTCTGTTTTGAGCTTATTTAGAGTCTAGTTTTCTCTACTTATA	3183	Db	4194	AGTCTGTGGCTTGACTGTTTTTCTTATGCTTTTGGATGAATAGAGTTTTTAAATTTTGAC	4253
QY	3004	AGTAGAAATGGTGAGATGTTTTTCTTACCTTAAGGGAGATGGTAAAGAACAT	3063	QY	4084	AAGGTCAAAATTTATTTTTTCTTTTGTGTGATATTTTTTCTCCAAATTTAAACCCCAAGA	4143
Db	3184	AGTAGAAATGGTGAGATGTTTTTCTTACCTTAAGGGAGATGGTAAAGAACAT	3243	Db	4254	AAGGTCAAAATTTATTTTTTCTTTTGTGTGATATTTTTTCTCCAAATTTAAACCCCAAGA	4313
QY	3064	GAATGTCTTTTTTCAAACTTTATTGACAAAGTGAATTTTCAAGTCTGTGTTTCAAAAATATAT	3123	QY	4144	TTTCAGATATTTCTGCTCTATTATATAAACTTTTATATATTTTATATTTTGTGATCTACCTGA	4203
Db	3244	GAATGTCTTTTTTCAAACTTTATTGACAAAGTGAATTTTCAAGTCTGTGTTTCAAAAATATAT	3303	Db	4314	TTTCAGATATTTCTGCTCTATTATATAAACTTTTATATATTTTATATTTTGTGATCTACCTGA	4373
QY	3124	TCATGACCTGTGATCCAGCAAGAGGATTCAGTCAAGAGTCACTCAACTGATTTAG	3183	QY	4204	ATTGATATGATGTTGTCGAATTTGATGATCAGGTTCTTTTTTCCCATACATAAGTATCC	4263
Db	3304	TCATGACCTGTGATCCAGCAAGAGGATTCAGTCAAGAGTCACTCAACTGATTTAG	3363	Db	4374	ATTGATATGATGTTGTCGAATTTGATGATCAGGTTCTTTTTTCCCATACATAAGTATCC	4433
QY	3184	TTGTTTTAGAGATGAGAAATGGAACTGAGGAAATGGAGGCAATTTTCCATGACTTCCC	3243	QY	4264	AGTCATTGTAAACACTGTTTTTATTTGAAGAAATTTATCTCTCATTTAAATTTACTTTGCCAA	4323
Db	3364	TTGTTTTAGAGATGAGAAATGGAACTGAGGAAATGGAGGCAATTTTCCATGACTTCCC	3423	Db	4434	AGTCATTGTAAACACTGTTTTTATTTGAAGAAATTTATCTCTCATTTAAATTTACTTTGCCAA	4493
QY	3244	TTGTTAAACAGAACACAGAGGCAAGAGGCTGGCTCTTACATCACTCTCACCTTCCA	3303	QY	4324	TTAGTAAATAATCAATTAACCATRMMRRR 4356	
Db	3424	TTGTTAAACAGAACACAGAGGCAAGAGGCTGGCTCTTACATCACTCTCACCTTCCA	3483	Db	4494	TTAGTAAATAATCAATTAACCATATAAAAAA 4526	
QY	3304	AATCTTGTGAAAGTGAATCTCTTCCAGAACCAAAATTAACCTTCTCAAGTCTGGCT	3363	RESULT 3			
Db	3484	AATCTTGTGAAAGTGAATCTCTTCCAGAACCAAAATTAACCTTCTCAAGTCTGGCT	3543	BC005293			
QY	3364	GCTTGAGGTGGAACCTCCAGCTCAAGGGAGTTAGGGAATGAGGCTTTTTTTTAAAG	3423	LOCUS			
Db	3544	GCTTGAGGTGGAACCTCCAGCTCAAGGGAGTTAGGGAATGAGGCTTTTTTTTAAAG	3603	DEFINITION			
QY	3424	CTTCTCAGCTTCTTAGGGAAACAGAAATTTGGGTGAGCCAACTCTGCAATTTCTACTACAGG	3483	BC005293			
Db	3604	CTTCTCAGCTTCTTAGGGAAACAGAAATTTGGGTGAGCCAACTCTGCAATTTCTACTACAGG	3663	ACCESSION			
QY	3484	CATTGAGACAGTTAGATTATGAAATATATATAGAGAGTTATGAACTTAAATTTATGAT	3543	BC005293.1			
Db	3664	CATTGAGACAGTTAGATTATGAAATATATATAGAGAGTTATGAACTTAAATTTATGAT	3723	GI:14710042			
QY	3544	AGTGGTATGACATTTGGATAGAACTGTTTATAGAGTATGAAATTTGACAGGCAATTT	3603	HTC.			
Db	3724	AGTGGTATGACATTTGGATAGAACTGTTTATAGAGTATGAAATTTGACAGGCAATTT	3783	Homo sapiens			
QY	3604	AGTTTGTAAATGGAGTCAATTTGAGTCTTTTAAATAGCCATGTATCATATTAACCAAGTGA	3663	Homo sapiens			
Db	3784	AGTTTGTAAATGGAGTCAATTTGAGTCTTTTAAATAGCCATGTATCATATTAACCAAGTGA	3843	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
QY	3664	AGCTGTGGAAATATATGTTCTCCATTTTACAGTTTAAAGGAATATAATTTGAGAGATTTAT	3723	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
Db	3844	AGCTGTGGAAATATATGTTCTCCATTTTACAGTTTAAAGGAATATAATTTGAGAGATTTAT	3903	1 (bases 1 to 1497)			

Strausberg, R.  
Direct Submission  
Submitted (27-MAR-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: CLONTECH Laboratories, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/INLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 16 Row: j Column: 6  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 3098674  
 This clone has the following problem: frame shifted.

## FEATURES

Location/Qualifiers  
 1..1497  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3997534"  
 /tissue\_type="Bladder, carcinoma"  
 /clone\_lib="NIH\_MGC\_53"  
 /lab\_host="DH10B"  
 /note="vector: pDNR-LIB"

## ORIGIN

Query Match 33.2%; Score 1460.2; DB 3; Length 1497;  
 Best Local Similarity 99.4%; Pred. No. 4.9e-302;  
 Matches 1476; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 22 ACCTGGCCCGCCAGCCTGAGAGCCGACCGATGTGGAACTGCTGCCGCCCGCGGCC 81  
 DB 2 ACCTGGCCCGCCAGCCTGAGAGCCGACCGATGTGGAACTGCTGCCGCCCGCGGCC 61

QY 82 GGCAGGAGGAGACCATACAGACTTTTGACTGGCTTGAGTAGCTGTGGAGGAAAA 141  
 DB 62 GGCAGGAGGAGACCATACAGACTTTTGACTGGCTTGAGTAGCTGTGGAGGAAAA 121

QY 142 CTGTGCCATTTCTAATTTGAAATGATCAGTCGATCAGCCGAAATCATGCTGTGTTAATCG 201  
 DB 122 CTGTGCCATTTCTAATTTGAAATGATCAGTCGATCAGCCGAAATCATGCTGTGTTAATCG 181

QY 202 TAACTTTTCTGTAAACCAACCTGAGTCAACAGATGAAATCCCTGTATTTGACATTAAGA 261  
 DB 182 TAACTTTTCTGTAAACCAACCTGAGTCAACAGATGAAATCCCTGTATTTGACATTAAGA 241

QY 262 TAATTTCTAGTATGTTACCTTTGTTAATGAGGAAAAATGCAATGGCTTTTCCCGAAC 321  
 DB 242 TAATTTCTAGTATGTTACCTTTGTTAATGAGGAAAAATGCAATGGCTTTTCCCGAAC 301

QY 322 TTTGAAGTCGGGGATGTTACTTTTGGAGTGTGGAGTAAATTCAGAAATAGAGTA 381  
 DB 302 TTTGAAGTCGGGGATGTTACTTTTGGAGTGTGGAGTAAATTCAGAAATAGAGTA 361

QY 382 TGAGCCTTTGGTTGCATGCTCTTTCTTTTGTAGATGCTCTGGGAAAACTGCTTTAAATCA 441  
 DB 362 TGAGCCTTTGGTTGCATGCTCTTTCTTTTGTAGATGCTCTGGGAAAACTGCTTTAAATCA 421

QY 442 AGCTATATTGCAACTGGAGGATTTACTGTAACAAATTTGGACAGAGATGCACTACCT 501  
 DB 422 AGCTATATTGCAACTGGAGGATTTACTGTAACAAATTTGGACAGAGATGCACTACCT 481

QY 502 TGTTCATGGTATCAGTGAAGTACCAATTAACAAATATGTGCACTATTTGTGGAGCTCC 561  
 DB 482 TGTTCATGGTATCAGTGAAGTACCAATTAACAAATATGTGCACTATTTGTGGAGCTCC 541

QY 562 AATTGTAAAGCCAGAAATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAAAGAGCAGCC 621  
 DB 542 AATTGTAAAGCCAGAAATTTTACTGAAATTCCTGAAAGCAGTTTCAGTCCAAAGAGCAGCC 601

QY 622 TCACAAATTTGAAGTTTATACCACCTCTTGTATGAACCATCTATTGGAAGTAAAAATGT 681  
 DB 602 TCACAAATTTGAAGTTTATACCACCTCTTGTATGAACCATCTATTGGAAGTAAAAATGT 661

QY 682 TGATCTGTACGAGCGCAGGAGAAACAAATCTTCAAGGGAACATTTATATTTT 741  
 DB 662 TGATCTGTACGAGCGCAGGAGAAACAAATCTTCAAGGGAACATTTATATTTT 721

QY 742 GAATGCCAAACAGCATAGAAATTTGATGTTCCGAGTTGCTTTTGGAGGTGGGAAAGCTAG 801  
 DB 722 GAATGCCAAACAGCATAGAAATTTGATGTTCCGAGTTGCTTTTGGAGGTGGGAAAGCTAG 781

QY 802 GTTGATAACAGAGAGAAATTCAGAGAAACATATTTCTTTTGGCTCCGCGAAGCTGTGT 861  
 DB 782 GTTGATAACAGAGAGAAATTCAGAGAAACATATTTCTTTTGGCTCCGCGAAGCTGTGT 841

QY 862 TGTGTATACAGGAATTAACAACTCAGAGACCTTAATTTCTGACTGTTCAGAGAAATGGAT 921  
 DB 842 TGTGTATACAGGAATTAACAACTCAGAGACCTTAATTTCTGACTGTTCAGAGAAATGGAT 901

QY 922 TCAGTCAATAATGGATATGCTCAAAAGGCAAGGCTTTAGACCTATTTCTTGAAGCAAAAT 981  
 DB 902 TCAGTCAATAATGGATATGCTCAAAAGGCAAGGCTTTAGACCTATTTCTTGAAGCAAAAT 961

QY 982 TGGATTGGCGGTGATTTTCATGACTACAAAGAAATTAATCTGTGATCCTCAGGGCCATCCCAG 1041  
 DB 962 TGGATTGGCGGTGATTTTCATGACTACAAAGAAATTAATCTGTGATCCTCAGGGCCATCCCAG 1021

QY 1042 TACAGGATTAAGAGCAACAACTCCAGGACCAAGCCCTTTTCAAGGGCTGTTCAGTTGATGA 1101  
 DB 1022 TACAGGATTAAGAGCAACAACTCCAGGACCAAGCCCTTTTCAAGGGCTGTTCAGTTGATGA 1081

QY 1102 AATACTAATGCCAAGCGCCCGAGTGAACACTCAACATACGCTAGCTGACACAGAAATCAGA 1161  
 DB 1082 AATACTAATGCCAAGCGCCCGAGTGAACACTCAACATACGCTAGCTGACACAGAAATCAGA 1141

QY 1162 GCAAGCAGATACATGGATTTGAGTGAAGGCGCAAGAAATCAAAAGTCTCCAAAATGGA 1221  
 DB 1142 GCAAGCAGATACATGGATTTGAGTGAAGGCGCAAGAAATCAAAAGTCTCCAAAATGGA 1201

QY 1222 ACACAAATTCAGAAATCTTTTCAAGACGACCCACTGTAAAGGAGTCTTCGCAAAACAAG 1281  
 DB 1202 ACACAAATTCAGAAATCTTTTCAAGACGACCCACTGTAAAGGAGTCTTCGCAAAACAAG 1261

QY 1282 CTCTAATAATAATAGTATGATCAATAATCTTTGGCTAAGATGAGAAATCCAAATATCA 1341  
 DB 1262 CTCTAATAATAATAGTATGATCAATAATCTTTGGCTAAGATGAGAAATCCAAATATCA 1321

QY 1342 GCTTTTCAACCACTTAATTTGCCAAGTATTAATAAGTAAGATAGGCTTCTCAGCAGCA 1401  
 DB 1322 GCTTTTCAACCACTTAATTTGCCAAGTATTAATAAGTAAGATAGGCTTCTCAGCAGCA 1381

QY 1402 GCAGACCAACTCCATCAGAAATCTTTTCCAGCGCTTACC-AAAAAAGGAAAGGAGTG 1460  
 DB 1382 GCAGACCAACTCCATCAGAAATCTTTTCCAGCGCTTACC-AAAAAAGGAAAGGAGTG 1441

QY 1461 AAGAAATCAAGAAATGCTTTCATGCAATTCAGCAAGATAGAAA 1505  
 DB 1442 AAGAAATCAAGAAATGCTTTCATGCAATTCAGCAAGATAGAAA 1486

## RESULT 4

AY416034 2265 bp DNA linear GSS 17-DBC-2003  
 LOCUS Homo sapiens NBS1 gene, VIRTUAL TRANSCRIPT, partial sequence,  
 DEFINITION genomic survey sequence.  
 AY416034  
 VERSION AY416034.1 GI:39771994  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Clark, A.G., Glanowski, S., Nielsen, R., Thomas, P., Kejariwal, A.,  
 Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, P., Murphy, B.,  
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Snieksy, J.J.,  
 Adams, M.D. and Cargill, M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 JOURNAL Science 302 (5652), 1960-1963 (2003)









KatoH,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Onisato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.  
 Direct Submission  
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL:http://genome.gsc.riken.jp/  
 URL:http://fantom.gsc.riken.jp/  
 Location/Qualifiers

## FEATURES

source

1..1731  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM DB:6330501B13"  
 /db\_xref="taxon:10090"  
 /clone="6330501B13"  
 /sex="male"  
 /tissue\_type="medulla oblongata"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="adult"  
 94..>1731  
 /note="unnamed protein product; nibrin (MGD|MG1:1351625, GBNM\_013752, evidence: BLASTN, 100%, match=1730) putative"  
 /codon\_start=1  
 /protein\_id="BAC27610.1"  
 /db\_xref="GI:26327733"  
 /translation="MWKLPAAGAPGEYRLLAGVYVVGKNGGLIENDOSIRN HAVLVNFTVSLSDTEIPTLIKDNKSKIFVNEEKMQTGLSTLTKGDRVTFVPSKRFVEPLVVCSSLDVSGKTVLNQILQGLTANNTEECVHLVMSAVKVIK TICALICGRPIIKPEYSEFLKAVESKQPPDIIESFPYDEPAIGSKSDVLSGRHER KOIFKGTVPFLNAQKHKLSSAVFGGGEARLMAEDDEEQSFSPAGTCVVDVGIIT NTQLIISHQSKWIHLIMDTLORNGLRPIPEABIGLAVIFMTTENVCPQOGPCTELK TTPGPSLSQVLSANGKIIIPSAFVNMVTVADTESEPADTCMPLSERPEVVKIPLGEO SRKLSQETFNIEAKPKPSSKANNVASDILVRGTFPSYQLSPMKPVPVANKNDWTSQQ QNSIKNYFPQTRKRDERDNEPLSSCKSSRLSELSLEQTQFAGPSLWKSREHQS QNATLDREADTSVSGMDIELNRKSPDRKPLPTLRP"

## CDS

## ORIGIN

Query Match 24.6%; Score 1083; DB 3; Length 1731;  
 Best Local Similarity 79.8%; Pred. No. 3.8e-221;  
 Matches 1320; Conservative 0; Mismatches 330; Indels 9; Gaps 3;  
 QY 37 CTGAGAGCGGACCGATGTGAAACTGCTGCCCGCGCGGCGCGGAGGAGGAAACC 96  
 DB 78 CTGAGAGCGGCGCCCATGTGGAAGCTGCTCCGCGCGCGGCTCCAGGAAACC 137  
 QY 97 ATACAGACTTTTGACGTGGCGGTGAGTACGTTGTTGGAAGAAAAAATGTGCCATTTAAT 156  
 DB 138 ATACCGACTTTTGGCGGCGGTGAGTACGTTGTTGGGAGGAAAAAATGTGGCATTTCTGAT 197  
 QY 157 TGAATATGATCAGTCGATCAGCGGAATCATGCTGTGTTAACTGCTAACTTTCTGTAC 216  
 DB 198 TGAATATGATCAGTCGATCAGTCGAAACCATGCTGCTTTAAACAGTAAACTTCCCTGTAAC 257  
 QY 217 CAACCTGAGTCAACAGATGAAATCCCTGTATTGACATTTAAAGATAAATTTCTAAGTATGG 276

DB 258 CAGTTTGTAGTCAAAACAGATGAAATTCCTACATTTAAACAATAAAGATAAATTTCTAAGTATGG 317  
 QY 277 TACCTTTGTTAATGAGGAAAAAATCGAGAAATGGCTTTTCCCGAACTTTGAAAGTCGGGGA 336  
 DB 318 AACCTTTGTTAATGAGGAAAAAATCGAGATGGTCTTCTCTGACGTTTGAAGACAGGAGA 377  
 QY 337 TGGTATTACTTTTGGAGTGTGTTGGAAGTAAATTTCAAGATAGAGTATGAGCCCTTTGGTTGC 396  
 DB 378 TAGAGTTACCTTTGGGGTGTGTTGAAAGTAAATTTCAAGATAGAGTATGAGCCCTTTGGTTGT 437  
 QY 397 ATGCTCTCTTTGTTTATAGTGTCTCTGGGAAAACTGCTTTTAAATCAAGCTATATTCAACT 456  
 DB 438 TTGTTCTTCTTGTGTTTATAGTGTCTCTGGGAAAACTGCTTTTAAATCAAGCTATATTCAAGCT 497  
 QY 457 TGGAGATTTACTGTAAACAATTTGACACAGAGAATGCACTCACCTTGTCTATGATATCAGT 516  
 DB 498 TGGAGGACTTACCGCAAAACAACTGGACGGAAGAATGTACCCACCTTGTCTGTGCGGAGT 557  
 QY 517 GAAAGTTACCATTTAAACAATATGTGCACTCATTTGTGAGCTGCCAATTTGTAAGCCAGA 576  
 DB 558 TAAAGTCAACATTAACAATATATGTGCACTCATTTGTGAGCTGCCAATTTAATAAACCAGA 617  
 QY 577 ATATTTTACTGAATTTCTGAAAGCAGTTCAGTCCAAGAACGAGCTCCCAAAATTTGAAG 636  
 DB 618 ATACTTTTCTGAAATTTCTCAAGCAGTTGAAATCTAAGAAAACAGCCTCCAGATATTTGAAAG 677  
 QY 637 TTTTACCACCTCTTGTATGAACCATTTATGGGAATAAAAATGTTGATCTGTCTCAGGAG 696  
 DB 678 TTTTACCACCTCTTGTATGAACCATTTATGGGAATAAAAATGTTGATCTGTCTCAGGAG 737  
 QY 697 GCAGAAAAAGAAAAAACAATCTTCAAGAGGAAAAACAATTTATATTTTGAATGCCAAAAACAGCA 756  
 DB 738 ACATGAAAGAAAAACAGATCTTCAAGAGGAAAAACAATTTGTTGTTTAAATGCCACAGCA 797  
 QY 757 TAAGAAATTTGAGTTCCGAGTTGCTTTTGGAGGTGGGAGAGCTAGGTTGATTAACAGAGA 816  
 DB 798 CAAGAAACTCAGCTCGGAGTTGCTTTTCGAGGTGGAGAACGAGCTGATGTCAGAGAAGA 857  
 QY 817 GAATGAAGAGAACATATTTCTTTTGGCTCCGGGAGAGCTGTGTTGTTGATACAGGAAT 876  
 DB 858 CGACGAAGAGGAACAGAGCTTTCTTTTTCAGCTCCCGGAAACCTGCGTGTGTTGATGTAGGAAT 917  
 QY 877 AACAACTCACAGACCTTAATTCCTGAGTGTGACAGAGAAATGGATTCAGTCAATATGA 936  
 DB 918 AACGATACAGCTCATTAATTTACACTCCAGAAAAAATGGATTCATTTGATTAATGA 977  
 QY 937 TATGCTCCAAAGGCAAGGTCTTAGACCTATTCTCTGAAGCAGAAAAATGGAATGGCGGTGAT 996  
 DB 978 TACACTTCAAGGAATGGTCTCAGACCTATTCTCTGAAGCGGAGATTGGATTGGCTGTTAT 1037  
 QY 997 TTTGATGACTACAAGAAATTAATCTGTGATCCTCAGGGCCATCCAGTACAGGATTAAGAC 1056  
 DB 1038 TTTTATGACTACAGAAATTAATCTGTAATCCGACGGCCAGCTTGTGACAGAAATTAAGAC 1097  
 QY 1057 AACAACTCCAGACCAAGCCTTTTCAAGGGGTGTGAGTTGATAAAAAATAATCCCAAG 1116  
 DB 1098 AACGACTCCAGGACCAAGCCTTTCCCAAGTCTGTGACCAATGGAAAAATATATCCCAAG 1157  
 QY 1117 GCGCCCGTGAACACTACAACATAGTGTGACACAGAAATCAGAGCAGAGATACATG 1176  
 DB 1158 TGCTCCAGTGAATATGACCAATAGTGTGACACAGAAATCAGAGCAGAGATACATG 1217  
 QY 1177 ---GGAATTTGAGTGAAGGCCAAAAGAAATCAAGTCTCCAAATGGGACAAAAATTCAG 1233  
 DB 1218 TATGCTTTGAGTGAAGACCAAGAAAGATTAAGATCCCTGGACTGGAAACAAAGCTCTAG 1277  
 QY 1234 AATGCTTTCAAGACGCAACCCACTGTATAAGAGGTCTTGCAGAAAAAAGCTCTTAATAATAA 1293  
 DB 1278 GAACTTTCAAGAGAAACATTCATATATAAGAGGAGCCCTTAAACCAAGCTCCAAAGCTAA 1337  
 QY 1294 TAGTATGATCAAAATACATTTTGGCTTAAGATGAGAAATCCCAAACTATACGCTTTTCAACCAAC 1353  
 DB 1338 CAACGTAGCATCAGATACGCTGGTTAGAGGAAAGACCCCGAGCTATACGCTTTTCTCCAAT 1397

QY 1354 TAAATTGCCAAGTATATAAATAAAGTAAAGATAGGGCTTCTCAGCAGCAGCAGCAACTC 1413  
 |||||  
 Db 1398 GAAATTTCTGTTGCAAAATAAATAAAGGATTGGACTTCTCAGCAGCAGCAG--AACTC 1454  
 |||||  
 QY 1414 CATCAGAACTACTTTTCAGCGCTTACCAAAAAAGGGAAGGATGAGAAATCAAGA 1473  
 |||||  
 Db 1455 CATCAAAACTACTTCCAGCATGCAACAGAAAAGGGAAGGATGAGAACCAACCGA 1514  
 |||||  
 QY 1474 AATGTCTTCATGCAAAATCAGCAAGATAGAAAGCTTGTGTCTTTTGAACAACACACA 1533  
 |||||  
 Db 1515 GCTGTCTCTGTCGAATCATCCAGGATGGAGCTGTCTGTCTCTTTTGAACAACACACA 1574  
 |||||  
 QY 1534 ACTGTACACCTCTATTTGGAAGAAATAAGGAGCAGCATCTATCTGAGAAATGAGCCTGT 1593  
 |||||  
 Db 1575 ACCTGCGGACCTCTACTGTGGAAGCAAGG--AGCATCAGTCTCAGAATGCGACCT 1631  
 |||||  
 QY 1594 GGACACAACTCAGCAATTAACCTATTATCAGATACAGATTAAATCTATTGTGAAGA 1653  
 |||||  
 Db 1632 GGACGGAAGCGACACCTCATCTGTGGGTGGATGGACATAGAACTCAACAGGAAGAG 1691  
 |||||  
 QY 1654 TTCTGCGAGTAAATCTCATGCTCGAGAAAGCTAAAGATC 1692  
 |||||  
 Db 1692 TCTGACAGGAACCACTTCCACAGAACTCTTAGACC 1730  
 |||||

RESULT 8  
 BX405940  
 LOCUS  
 DEFINITION BX405940 Homo sapiens FETAL LIVER Homo sapiens cDNA clone  
 CS0DM009Y012 5-PRIME, mRNA sequence.  
 ACCESSION BX405940  
 VERSION  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 964)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 15, 2003 this sequence version replaced gi:30766648.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 This sequence belongs to sequence cluster 4762.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?sc=CS0AM009BH06QPl&c=4762.r.

FEATURES  
 source  
 1..964  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DM009Y012"  
 /tissue\_type="FETAL LIVER"  
 /dev\_stage="fetal"  
 /clone\_lib="Homo sapiens FETAL LIVER"  
 /notes="Organ: liver; Vector: pCMVSPORT\_6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-strand cDNA was digested with Not I and  
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
 vector. Library was not normalized."

## ORIGIN

Query Match 20.7%; Score 913.2; DB 5; Length 964;  
 Best Local Similarity 94.1%; Pred. No. 8.8e-185;

RESULT 9  
 BX405939/c

Matches 907; Conservative 34; Mismatches 23; Indels 0; Gaps 0;  
 QY 1042 TACAGATTAAAGACAAACAATCCAGGACCAAGCCCTTTTCAAGGGCGTGTCACTGATGA 1101  
 |||||  
 Db 1 TACAGATTAAAGACAAACAATCCAGGACCAAGCCCTTTTCAAGGGCGTGTCACTGATGA 60  
 |||||  
 QY 1102 AAACTAATGCCNAGGGCCCCAGTGAACACTCAACATATAGTGTGACACAGAATCAGA 1161  
 |||||  
 Db 61 AAAACTAATGCCNAGGGCCCCAGTGAACACTCAACATATAGTGTGACACAGAATCAGA 120  
 |||||  
 QY 1162 GCAAGCAGATACATGGGATTTTGAGTGAAAGGCCAAAGAAATCAAAAGTCTCCAAAATGGA 1221  
 |||||  
 Db 121 GCAAGCAGATACATGGGATTTTGAGTGAAAGGCCAAAGAAATCAAAAGTCTCCAAAATGGA 180  
 |||||  
 QY 1222 ACAAAATTTCAGAAATGCTTTTCAACAGCAGCCACTGTAAAGGAGTCTCTGCAAAACAAG 1281  
 |||||  
 Db 181 ACAAAATTTCAGAAATGCTTTTCAACAGCAGCCACTGTAAAGGAGTCTCTGCAAAACAAG 240  
 |||||  
 QY 1282 CTCTAATAATATAGTATGGTATCAAAATCTTTGGCTAAGATGAGAAATCCAAAATCTATCA 1341  
 |||||  
 Db 241 CTCTAATAATATAGTATGGTATCAAAATCTTTGGCTAAGATGAGAAATCCAAAATCTATCA 300  
 |||||  
 QY 1342 GCTTTTCAACCAACTAAATTCGCCAAGTATAATAAAGTAAAGATAGGCTTCTCAGCAGCA 1401  
 |||||  
 Db 301 GCTTTTCAACCAACTAAATTCGCCAAGTATAATAAAGTAAAGATAGGCTTCTCAGCAGCA 360  
 |||||  
 QY 1402 GCAGACCAACTCCATCAGAAACTTACTTTTCAGCGCTCTACCAAAAAAGGAAAGGATGA 1461  
 |||||  
 Db 361 GCAGACCAACTCCATCAGAAACTTACTTTTCAGCGCTCTACCAAAAAAGGAAAGGATGA 420  
 |||||  
 QY 1462 AGAAATCAAGAAATGTCTTCATGCAAAATCAGCAAGATAGAAACCTCTTGTCTCTTTT 1521  
 |||||  
 Db 421 AGAAGATCAAGAGATGTCTTCATGCGAATCAGCRAGATAGRAGCGTCTTGTCTCTTTT 480  
 |||||  
 QY 1522 AGAACAACAACACCTGCTACACCTCTATTGTGGAAAAATAGGAGCAGCATCTATCTGA 1581  
 |||||  
 Db 481 AGAACAACAACACCTGCTACACCTCTATTGTGGAAAGGTGAGGAGCAGCATCTATCTGA 540  
 |||||  
 QY 1582 GAATGAGCCTGTGGACACAAACTCAGACAAATACTTATTACAGATACAGATTAAATC 1641  
 |||||  
 Db 541 GAATGAGCCTGTGGACACAAACTCAGACAAATACTTATTACAGATACAGATTAAATC 600  
 |||||  
 QY 1642 TATTGTGAAAAATTTCTGCCAGTAAATCTCATGCTGCAGAAAAAGCTTAAGATCAAAATAAA 1701  
 |||||  
 Db 601 TATTGTGAAAAATTTCTGCCAGTAAATCTCATGCTGCAGAAAAAGCTTAAGATCAAAATAAA 660  
 |||||  
 QY 1702 AAGGGAATGGATGATGTGGCCATAGAGATGAGTATTGGAAACAGTATTCAAGGACAC 1761  
 |||||  
 Db 661 GAGGGAATGGATGATGTGGCCATAGAGATGAGTATTGGAAACAGTATTCAAGGACAC 720  
 |||||  
 QY 1762 AAAACAGAGTTAGAAATTTGATGTGAAAGTTCAAAAACAGGAGGAGATGTCAATGTTAG 1821  
 |||||  
 Db 721 AAAACAGAGTTAGAAATTTGATGTGAAAGTTCAAAAACAGGAGGAGATGTCAATGTTAG 780  
 |||||  
 QY 1822 AAAAAGGCCAAGATGGATATAGAAACAAATGACACTTTTCAGTGATGAAGCAGTACCAGA 1881  
 |||||  
 Db 781 AARAGGCCAAGATGGATATAGAAACAAATGACACTTTTCAGTGATGAAGCAGTACCAGA 840  
 |||||  
 QY 1882 AAGTAGCAAAATATCTCAAGAAAAATGAATTTGGGAAGAACCTGAACTCAAGGAGACTC 1941  
 |||||  
 Db 841 AAGTAGCAAAATATCTCAAGAAAAATGAATTTGGGAAGAACCTGAACTCAAGGAGACTC 900  
 |||||  
 QY 1942 ACTATGGTCAAGTAAAGAAATATCTAAACAATGACAACTTCAGGATGATAGTGGATGCT 2001  
 |||||  
 Db 901 ACTATGGTCAAGTAAAGAAATATCTAAACAATGACAACTTCAGGATGATAGTGGATGCT 960  
 |||||  
 QY 2002 TCCA 2005  
 |||||  
 Db 961 TCCA 964  
 |||||





```
Db 201 AGGAAAAATTATATTTTGAATGCCAAACAGCATAGAATTCGATCCGACGTGT 260
QY 781 CTTTGGAGTGGGGAAGCTAGGTTGATATACAGAGAAGATGAAGAAGACATAATTTCTT 840
Db 261 CTTTGGAGTGGGGAAGCTAGGTTGATATACAGAGAAGATGAAGAAGACATAATTTCTT 320
QY 841 TTTGGCTCCGGGAACGTGTGTTGATACAGGAATAACAACTCAGACCTTAATTC 900
Db 321 TTTGGCTCCGGGAACGTGTGTTGATACAGGAATAACAACTCAGACCTTAATTC 380
QY 901 TGAATCTCAGAGAAATGGATTGAGTCAATATGATATGCTCCAAAGGCAAGGCTTAG 960
Db 381 TGAATCTCAGAGAAATGGATTGAGTCAATATGATATGCTCCAAAGGCAAGGCTTAG 440
QY 961 ACCTATTCTCGAAGCAAGAAATGGATTGAGTGGCGGTGATTTTCATGACTACAAAGAAATTCG 1020
Db 441 ACCTATTCTCGAAGCAAGAAATGGATTGAGTGGCGGTGATTTTCATGACTACAAAGAAATTCG 500
QY 1021 TGAATCTCAGAGCAATCCAGTACAGGATTAAGACAACAACTCCAGGACCAAGCCTTTC 1080
Db 501 TGAATCTCAGAGCAATCCAGTACAGGATTAAGACAACAACTCCAGGACCAAGCCTTTC 560
QY 1081 ACAAGGCGTGTGATGATGAAAAAATAATGCCAAGCGCCCAAGTGAACTACACATA 1140
Db 561 ACNAGGCGTGTGATGATGAAAAAATAATGCCAAGCGCCCAAGTGAACTACACATA 620
QY 1141 CGTAGCTGACACAGAATCAGAGCAAGCAGATACATGGGATTTGAGTGAAGGCCAAGAA 1200
Db 621 CGTAGCTGACACAGAATCAGAGCAAGCAGATACATGGGATTTGAGTGAAGGCCAAGAA 680
QY 1201 AATCAAGTCTCCAAATCGMACAAAAATTCAGAAATGCTTTTCCAGAGCGCACCTGCT 1260
Db 681 AATCAAGTCTCCAAATCGMACAAAAATTCAGAAATGCTTTTCCAGAGCGCACCTGCT 740
QY 1261 AAAGGAGTCTGCAAAACCAAGCTCTAATAATAGTATGGTATCAAAATCTTTGGCTAA 1320
Db 741 NAAGGAGTCTGCANAAATAGCTCTAAT-ATAATAGTATGGGATCAAAATCTCTGCTAA 799
QY 1321 GATGAGAAATCCCAACTATCAGCTTTTCCAACTAAATTTGCCAAGTATATAATAAAGTAA 1380
Db 800 GATGAGAAATCCCAACTATCAGCTTTTCCAACTAAATTTGCCAAGTATATAATAAAGTAA 859
QY 1381 AGATAGGCTCTCAGCAGCAGCAGACCAACTCCATCAGAAATCTTTTCAGCGCTAC 1440
Db 860 AGATAGGCTCTCAGCAGCAGCAGACCAACTCCATCAGAAATCTTTTCAGCGCTAC 919
QY 1441 CAAAAA 1447
Db 920 CAAAAA 926

RESULT 12
LOCUS AY416036
DEFINITION Mus musculus NBS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
GENOMIC SURVEY SEQUENCE.
ACCESSION AY416036
VERSION AY416036.1 GI:39771996
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1803)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
```

```
REFERENCE 2 (bases 1 to 1803)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1803
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1803
/gene="NBS1"
/locus_tag="HCM5756"

ORIGIN
Query Match 18.7%; Score 822.2; DB 9; Length 1803;
Best Local Similarity 71.2%; Pred. No. 3.1e-165;
Matches 988; Conservative 0; Mismatches 393; Indels 6; Gaps 2;

QY 224 AGTCAACAGATGAAATCCCTGATTGACATTAAAGAGATAATCTTAAGTATGTTACCTTT 283
Db 1 AGTCAACAGATGAAATCCCTGATTGACATTAAAGAGATAATCTTAAGTATGTTACCTTT 60
QY 284 GTTAATGAGGAAAAAATCGAATCGCTTTTCCCGAACTTTGAAAGTCGGGGATGGTATT 343
Db 61 GTTAATGAGGAAAAAATCGAATCGCTTTTCCCGAACTTTGAAAGTCGGGGATGGTATT 120
QY 344 ACTTTTGAGTGTTTGGAGTAAATTCAGAATAGATGATGAGCCTTTGGTTCATGCTCT 403
Db 121 ACTTTTGAGTGTTTGGAGTAAATTCAGAATAGATGATGAGCCTTTGGTTCATGCTCT 180
QY 404 TCTTGTGTAGATGTCTCTGGGAAAACTGCTTTAAATCAAGCTATATTGCAACTGGAGGA 463
Db 181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
QY 464 TTTACTGTAAACAATTGGACAGAAATGCACTCACTTGTGTCATGGTATCAGTGAAGTT 523
Db 241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
QY 524 ACCATTAAACAAATATGTGCACTCATTTGTGGACGCTCCAAATTTGTAAGCCAGAAATTTT 583
Db 301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360
QY 584 ACTGAAATCTCGAAAGCAGTTTCAGTCCCAAGAGCAGCCTCCCAAAATGAAAGTTTTCAC 643
Db 361 TCTGAAATCTCGAAAGCAGTTTCAGTCCCAAGAGCAGCCTCCCAAAATGAAAGTTTTCAC 420
QY 644 CCACCTCTTGATGAACCATCTATTGGAAGTAAAAATGTTGATCTGTCTAGGACGGCAGGAA 703
Db 421 CCACCATCTGATGAACCATCTATTGGAAGTAAAAATGTTGATCTGTCTAGGACGGCAGGAA 480
QY 704 AGAAAAACAATCTTCAAGGGGAAAAATTTATATTTTGAATGCCCAACAGCATGAAGAA 763
Db 481 AGAAAAACAATCTTCAAGGGGAAAAATTTATATTTTGAATGCCCAACAGCATGAAGAA 540
QY 764 TTGAGTTCCGCAAGTTGTCTTTGGAGGTGGGAGCTAGGTTTGATTAACAGAGAGTAATGAA 823
Db 541 CTCAGCTCGGCAAGTTGCTTTTCGAGGTGGAGAGCCAGGCTGATGGCAGAGCAGCAGAA 600
QY 824 GAAGAACATAATTTCTTTTGGCTCCGGGAACGTTGTTGTTGATACAGGAATAACAAC 883
Db 601 GAGGAACAGAGCTTCTTTTCAGCTCCCGGAACCTCGTGTGTTGATGTAGGAATAACGAT 660
QY 884 TCACAGACTTAATTTCTGCTGCTGCTCAGAGAAATGATTCAGTCAATATGATATGCTC 943
Db 661 ACACAGCTCATAAATTTTCACACTCCCAAGAAAAATGATTCATTTGATATGATACACT 720
QY 944 CAAAGGCAAGGCTTTTAGACCTATTCTCTGAAAGCAGAAATTTGGATTTGGCGTGTATTTTCATG 1003
```



[illegible]

FEATURES  
source

Location/Qualifiers  
1. .789  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/cell\_line="HT1080"  
/clone\_lib="Athersys RAGE Library"  
/notes="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 16.8%; Score 741; DB 4; Length 789;  
Best Local Similarity 97.8%; Pred. No. 7.1e-148;  
Matches 772; Conservative 0; Mismatches 15; Indels 2; Gaps 2;  
  
QY 637 TTTTACCCACCTCTTGATGAAC-CATCTATTGGAAGTAAAAATGTTGATCTCTCAGGAC 695  
DB |||||  
2 TTTTTTACCACCTCTTGATGAACATCATCTATTGGAAGTAAAAATGTTGATCTCTCAGGAC 61  
  
QY 696 GGCAGGAAAGAAACAAATCTTCAAAGGGNAACATTATATTTTGAATGCCAAACAGC 755  
DB |||||  
62 GGCAGGAAAGAAACAAATCTTCAAAGGGNAACATTATATTTTGAATGCCAAACAGC 121  
  
QY 756 ATAAGAAATTTGAGTTCGGCAGTTGTCTTTGGAGGTGGGGAAGCTAGCTTGATTAACAGNAG 815  
DB |||||  
122 ATAGAAATTTGAGTTCGGCAGTTGTCTTTGGAGGTGGGGAAGCTAGCTTGATTAACAGNAG 181  
  
QY 816 AGAATGAAGAAGAACATAATTTCTTTTGGCTCCGGGAACGTCGTGTTGTTGATACAGGAA 875  
DB |||||  
182 AGAATGAAGAAGAACATAATTTCTTTTGGCTCCGGGAACGTCGTGTTGTTGATACAGGAA 241  
  
QY 876 TAACAACCTCAGAGACCTTAATTTCTCTGACTGTCAGAGAATAATGGATTCAGTCAATAATGG 935  
DB |||||  
242 TAACAACCTCAGAGACC-TGATTCCTGACTGTCAGAGAATAATGGATTCAGTCAATAATGG 300  
  
QY 936 ATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTGAAGCAGAAATTTGGATTGGCGGTGA 995  
DB |||||  
301 ATATGCTCCAAAGGCAAGGTCTTAGACCTATTCTGAAGCAGAAATTTGGATTGGCGGTGA 360  
  
QY 996 TTTTTCATGACTACAAAGAAATTTACTGTGATCTCTCAGGGCCATCCAGTACAGGATTAAGA 1055  
DB |||||  
361 TTTTTCATGACTACAAAGAAATTTACTGTGATCTCTCAGGGCCATCCAGTACAGGATTAAGA 420  
  
QY 1056 CAACAACCTCAGAGACCAAGCCTTTCAAGGGCGTGTGATGATGAAAACTAATGCCAA 1115  
DB |||||  
421 CAACAACCTCAGAGACCAAGCCTTTCAAGGGCGTGTGATGATGAAAACTAATGCCAA 480  
  
QY 1116 GCGCCCGAGTGAACACTTACAACATAGCTAGCTGACAGAACTCAGAGCAAGCAGATACAT 1175  
DB |||||  
481 GCGCCCGAGTGAACACTTACAACATAGCTAGCTGACAGAACTCAGAGCAAGCAGATACAT 540  
  
QY 1176 GGGATTTGAGTGAAGGCCCAAAAGAAATCAAAGTCTCCAAAATGGAACAAAATTCAGAA 1235  
DB |||||  
541 GGGATTTGAGTGAAGGCCCAAAAGAAATCAAAGTCTCCAAAATGGAACAAAATTCAGAA 600  
  
QY 1236 TGCTTTTCAAGACGACCCACTGTAAAGAGTCTTGCAAAACAAAGCTCTAATAATAATA 1295  
DB |||||  
601 TGCTTTTCAAGACGACCCACTGTAAAGAGTCTTGCAAAACAAAGCTCTAATAATAATA 660  
  
QY 1296 GTATGATCAATAACTTTTGGCTTAAGATGAGATCCCAACTATCATGCTTTCACCAACTA 1355  
DB |||||  
661 GCATGATCAATAACTTTTGGCTTAAGATGAGATCCCAACTATCATGCTTTCACCAACTA 720  
  
QY 1356 AATTGCCAAGTATAAATAAGATAAGATAGGCTTTCTCAGCAGCAGCAGACCAACTCCA 1415  
DB |||||  
721 AATTGCCAAGTATAAATAAGATAAGATAGGCTTTCTCATCAGCAGCAGACCAACTCCA 780  
  
QY 1416 TCAGAAACT 1424  
|||

Db 781 TCAGAGACT 789

Search completed: August 30, 2005, 14:45:13  
Job time : 13123 secs

**THIS PAGE BLANK (USPTO)**